



UTPL
La Universidad Católica de Loja

Vicerrectorado de Modalidad Abierta y a Distancia

Método Científico y Análisis de Datos

Guía didáctica





Facultad Ciencias Exactas y Naturales

Método Científico y Análisis de Datos

Guía didáctica

Carrera

PAO Nivel

Gestión Ambiental

III

Autor:

Diego Paúl Vélez Mora



Método Científico y Análisis de Datos

Guía didáctica

Diego Paúl Vélez Mora

Diagramación y diseño digital

Ediloja Cía. Ltda.

Marcelino Champagnat s/n y París

edilojacialtda@ediloja.com.ec

www.ediloja.com.ec

ISBN digital -978-9942-47-022-5

Año de edición: abril, 2024

Edición: primera edición reestructurada en enero 2025 (con un cambio del 30%)

Loja-Ecuador



**Reconocimiento-NoComercial-Compartirlgual
4.0 Internacional (CC BY-NC-SA 4.0)**

Usted acepta y acuerda estar obligado por los términos y condiciones de esta Licencia, por lo que, si existe el incumplimiento de algunas de estas condiciones, no se autoriza el uso de ningún contenido.

Los contenidos de este trabajo están sujetos a una licencia internacional Creative Commons **Reconocimiento-NoComercial-Compartirlgual 4.0 (CC BY-NC-SA 4.0)**. Usted es libre de **Compartir – copiar y redistribuir el material en cualquier medio o formato. Adaptar – remezclar, transformar y construir a partir del material citando la fuente, bajo los siguientes términos: Reconocimiento- debe dar crédito de manera adecuada, brindar un enlace a la licencia, e indicar si se han realizado cambios.** Puede hacerlo en cualquier forma razonable, pero no de forma tal que sugiera que usted o su uso tienen el apoyo de la licenciante. **No Comercial-no puede hacer uso del material con propósitos comerciales. Compartir igual-Si remezcla, transforma o crea a partir del material, debe distribuir su contribución bajo la misma licencia del original.** No puede aplicar términos legales ni medidas tecnológicas que restrinjan legalmente a otras a hacer cualquier uso permitido por la licencia. <https://creativecommons.org/licenses/by-nc-sa/4.0/>

Índice

1. Datos de información	9
1.1 Presentación de la asignatura.....	9
1.2 Competencias genéricas de la UTPL.....	9
1.3 Competencias del perfil profesional	9
1.4 Problemática que aborda la asignatura	10
2. Metodología de aprendizaje	11
3. Orientaciones didácticas por resultados de aprendizaje.....	12
Primer bimestre	12
Resultado de aprendizaje 1:	12
Contenidos, recursos y actividades de aprendizaje recomendadas.....	12
Semana 1	12
Unidad 1. El problema de investigación.....	13
1.1. ¿Qué es investigación y cómo comenzamos a investigar?.....	13
1.2. El planteamiento del problema de investigación cuantitativa	16
1.3. Componentes del planteamiento de investigación cuantitativa.....	19
Actividades de aprendizaje recomendadas	22
Autoevaluación 1	22
Resultado de aprendizaje 2:	26
Contenidos, recursos y actividades de aprendizaje recomendadas.....	26
Semana 2	26
Unidad 2. El marco teórico	26
2.1. Elaboración del marco teórico	26
2.2. Fases para la elaboración del marco teórico	28
Actividad de aprendizaje recomendada	33
Autoevaluación 2.....	34
Resultado de aprendizaje 3:	37
Contenidos, recursos y actividades de aprendizaje recomendadas.....	37
Semana 3	37

Unidad 3. Los alcances en la investigación cuantitativa	37
3.1. Alcance exploratorio	39
3.2. Alcance descriptivo.....	40
3.3. Alcance correlacional	40
3.4. Alcance explicativo	41
Actividades de aprendizaje recomendadas	41
Autoevaluación 3.....	42
Contenidos, recursos y actividades de aprendizaje recomendadas.....	44
Semana 4	44
Unidad 4. Formulación de hipótesis y diseños de investigación	44
4.1. Formulación de hipótesis de acuerdo con el alcance del estudio.....	44
4.2. Tipos de diseño en la investigación cuantitativa	47
Actividades de aprendizaje recomendadas	51
Autoevaluación 4.....	52
Resultado de aprendizaje 4:	55
Contenidos, recursos y actividades de aprendizaje recomendadas.....	55
Semana 5	55
Unidad 5. Introducción a la estadística descriptiva y tipos de variables....	56
5.1. Introducción a la estadística descriptiva.....	56
5.2. Población y muestra	60
5.3. Variables y datos.....	61
5.4. Tipos de variables	61
Actividad de aprendizaje recomendada	63
Contenidos, recursos y actividades de aprendizaje recomendadas.....	64
Semana 6	64
Unidad 5. Introducción a la estadística descriptiva y tipos de variables....	64
5.5. Descripción de datos por medio de gráficas.....	66
Actividad de aprendizaje recomendada	74
Autoevaluación 5.....	74

Contenidos, recursos y actividades de aprendizaje recomendadas.....	75
Semana 7.....	75
Unidad 6. Descripción de datos con medidas numéricas	76
6.1. Descripción de un conjunto de datos con medidas numéricas	78
6.2. Medidas de centro	80
6.3. Medidas de variabilidad.....	80
6.4. Mediciones de posición relativa.....	82
6.5. El resumen de cinco números y la gráfica de caja.....	85
Actividades de aprendizaje recomendadas	86
Autoevaluación 6.....	87
Resultado de aprendizaje 1 a 4:.....	90
Contenidos, recursos y actividades de aprendizaje recomendadas.....	90
Semana 8.....	90
Actividades de aprendizaje recomendadas	90
Segundo bimestre.....	93
Resultado de aprendizaje 4:	93
Contenidos, recursos y actividades de aprendizaje recomendadas.....	93
Semana 9.....	93
Unidad 7. Descripción de datos bivariados	93
7.1. Datos bivariados	96
7.2. Gráficas para variables cualitativas	96
7.3. Gráficas de dispersión para dos variables cuantitativas.....	98
7.4. Medidas numéricas para datos cuantitativos bivariados	100
Actividades de aprendizaje recomendadas	101
Autoevaluación 7.....	102
Resultado de aprendizaje 5:	105
Contenidos, recursos y actividades de aprendizaje recomendadas.....	105
Semana 10.....	105
Unidad 8. Probabilidad y distribuciones de probabilidad	105

8.1. El papel de la probabilidad en estadística	105
8.2. Eventos y el espacio muestral.....	110
8.3. Cálculo de probabilidades con el uso de eventos sencillos.....	111
Actividad de aprendizaje recomendada	111
Contenidos, recursos y actividades de aprendizaje recomendadas.....	112
Semana 11	112
Unidad 8. Probabilidad y distribuciones de probabilidad	112
8.4. Relaciones de evento y reglas de probabilidad.....	112
8.5. Independencia, probabilidad condicional y la regla de la multiplicación	114
8.6. Variables aleatorias discretas y sus distribuciones de probabilidad	114
Actividades de aprendizaje recomendadas	117
Autoevaluación 8.....	118
Contenidos, recursos y actividades de aprendizaje recomendadas.....	120
Semana 12	120
Unidad 9. Algunas distribuciones discretas útiles	120
9.1. La distribución binomial de probabilidad	122
9.2. La distribución de probabilidad de Poisson	126
Actividad de aprendizaje recomendada	128
Autoevaluación 9.....	128
Contenidos, recursos y actividades de aprendizaje recomendadas.....	130
Semana 13	130
Unidad 10. La distribución normal de probabilidad	130
10.1. Distribuciones de probabilidad para variables aleatorias continuas	133
10.2. La distribución normal de probabilidad.....	134
10.3. Áreas tabuladas de la distribución normal de probabilidad.....	136
Actividad de aprendizaje recomendada	137
Autoevaluación 10.....	138

Contenidos, recursos y actividades de aprendizaje recomendadas.....	140
Semana 14.....	140
Unidad 11. Distribuciones muestrales	140
11.1. Planes muestrales y diseños experimentales.....	144
11.2. Estadística y distribuciones muestrales	145
11.3. El teorema del límite central.....	146
11.4. La distribución muestral de la media muestral.....	147
Actividad de aprendizaje recomendada	149
Autoevaluación 11.....	149
Resultado de aprendizaje 6:	152
Contenidos, recursos y actividades de aprendizaje recomendadas.....	152
Semana 15.....	152
Unidad 12. El reporte de investigación	152
12.1. Reporte de resultados del proceso de investigación	155
Actividades de aprendizaje recomendadas	157
Autoevaluación 12.....	159
Resultado de aprendizaje 4 a 6:.....	162
Contenidos, recursos y actividades de aprendizaje recomendadas.....	162
Semana 16.....	162
Actividades finales del bimestre	162
4. Autoevaluaciones	165
5. Referencias bibliográficas	195



1. Datos de información

1.1 Presentación de la asignatura



1.2 Competencias genéricas de la UTPL

- Orientación a la innovación y a la investigación.
- Comunicación oral y escrita.
- Organización y planificación del tiempo.

1.3 Competencias del perfil profesional

- Conoce y aplica los pasos para el diseño de investigación.
- Realiza revisiones bibliográficas efectivas.
- Plantea hipótesis en torno al tema de investigación.
- Levanta, analiza e interpreta información de campo.
- Genera investigaciones científicas y aplica sus resultados como aporte a la solución de problemas ambientales.
- Escribe reportes de investigación siguiendo las normas técnicas y científicas.

1.4 Problemática que aborda la asignatura

Débil sustento técnico-científico a las propuestas de manejo y conservación de los recursos ambientales.





2. Metodología de aprendizaje

Para adquirir los conocimientos en esta materia, se empleará la metodología de aprendizaje **basado en análisis de estudio de caso**.

Esta metodología se utiliza en situaciones del mundo real como base para el aprendizaje y la aplicación de conceptos. Con este enfoque, se sumergirá en situaciones específicas, enfrentando problemas o desafíos que reflejan situaciones reales. Estos casos suelen presentar contextos complejos, permitiéndole explorar, analizar, comprender y resolver problemas de manera activa al finalizar cada semana de estudio.

Por ejemplo, podrá analizar un caso donde una organización ambiental desea evaluar la efectividad de su programa de conservación, centrado en la protección de una especie en peligro de extinción. Se le ofrece una introducción teórica sobre los diferentes diseños de investigación, revisa los elementos necesarios para elaborar un objetivo de investigación, desarrolla el objetivo de acuerdo con lo que comprendió y finalmente revisa la retroalimentación del profesor para verificar cuáles fueron sus aciertos y desaciertos en la elaboración de objetivos, fortaleciendo la comprensión de los contenidos.





3. Orientaciones didácticas por resultados de aprendizaje



Primer bimestre

Resultado de aprendizaje 1:

Conoce y aplica los pasos para el diseño de investigación.

La finalidad del resultado de aprendizaje es comprender los pasos esenciales para el planteamiento del problema de investigación. Para esto, comenzaremos con una revisión teórica de los conceptos acerca de la investigación cuantitativa, investigación cualitativa y el problema de investigación.

Contenidos, recursos y actividades de aprendizaje recomendadas

Recuerde revisar de manera paralela los contenidos con las actividades de aprendizaje recomendadas y actividades de aprendizaje evaluadas.



Semana 1

Estimado estudiante, bienvenido a la asignatura *Método Científico y Análisis de Datos*. En esta asignatura, exploraremos los principios del método científico, desde la formulación de objetivos y preguntas de investigación hasta la interpretación de resultados. Además, nos sumergiremos en las técnicas y herramientas esenciales para explorar y analizar datos, que le permitirán adquirir habilidades para la investigación y la toma de decisiones en el campo de la gestión ambiental.



En las unidades 1 a 5 abordaremos los fundamentos de la metodología de investigación, mientras que en las unidades posteriores nos enfocaremos en probabilidad y estadística.

Unidad 1. El problema de investigación

1.1. ¿Qué es investigación y cómo comenzamos a investigar?

La investigación desempeña un papel fundamental en el avance del conocimiento en todas las disciplinas, y los enfoques de investigación cuantitativa y cualitativa, se destacan como herramientas esenciales para abordar preguntas de investigación de manera integral.

Para comprender los fundamentos de la investigación científica, analizaremos los siguientes conceptos clave:

- Conceptos fundamentales de la investigación: la investigación es un conjunto sistemático de procesos que busca generar conocimiento mediante la aplicación del método científico. Se caracteriza por ser sistemática, controlada, empírica y crítica.
- Proceso inicial de investigación: la investigación comienza con una idea o problema que requiere resolverse. Esta idea debe refinarse y estructurarse formalmente para convertirse en un planteamiento más preciso del problema a investigar.
- Características del enfoque cuantitativo: sigue una secuencia rigurosa, utiliza la recolección de datos numéricos para probar hipótesis mediante análisis estadístico. Sus resultados buscan ser generalizables y replicables.
- Características del enfoque cualitativo: se enfoca en comprender fenómenos desde la perspectiva de los participantes en su ambiente natural. Es más flexible en su proceso y evoluciona con el desarrollo del estudio.
- Diferencias entre investigación cuantitativa y cualitativa: La investigación cuantitativa mide fenómenos, utiliza estadística y prueba hipótesis, mientras que la cualitativa explora los fenómenos en profundidad, se conduce en ambientes naturales y los significados se extraen de los datos.

Estas dos rutas de investigación no son mutuamente excluyentes; de hecho, pueden complementarse para lograr una comprensión más completa del fenómeno estudiado. La elección del enfoque dependerá del problema de investigación y los objetivos del estudio (Munch Galindo, 2009; Hernández-Sampieri, 2018; Domínguez & Tostado, 2016)

A raíz de lo mencionado anteriormente, responda estas preguntas para evaluar su comprensión acerca de este tema:

- ¿Cuáles serían las diferencias entre la investigación cuantitativa y cualitativa?
- ¿Cree que la investigación cuantitativa es mejor que la investigación cualitativa? Relacione estos conceptos con sus propias experiencias o situaciones del mundo real.

Nota: por favor, responda las preguntas en un cuaderno o documento Word.

Estas preguntas le ayudarán a evaluar su comprensión y a identificar áreas que requieren más atención. Esto facilitará la retención y aplicación del conocimiento.

La *investigación cuantitativa* se caracteriza por la recopilación y análisis de datos numéricos para entender patrones, relaciones y tendencias.

Este enfoque se basa en la objetividad, la medición y la estadística para proporcionar respuestas cuantificables a preguntas de investigación específicas. Si se sigue rigurosamente el proceso de este enfoque y, de acuerdo con ciertas reglas lógicas, los datos generados poseen los estándares de validez y confiabilidad, las conclusiones derivadas contribuirán a la generación de conocimiento. Por el contrario, la *investigación cualitativa*, adopta un enfoque más global y exploratorio. Se centra en la comprensión profunda de fenómenos sociales y humanos a través de métodos como entrevistas, observación participante y análisis de contenido. El proceso de este enfoque es más “circular” y dinámico, en lo cual la muestra, la recolección y el análisis son fases que se realizan prácticamente de manera simultánea.

Aunque distintos en su naturaleza, la investigación cuantitativa y cualitativa no son mutuamente excluyentes. En muchos casos, la combinación de estos enfoques a través de la *investigación mixta* fortalece la validez y la aplicabilidad de los hallazgos. El estudiante será desafiado a integrar estos métodos de manera reflexiva, seleccionando el enfoque más apropiado para abordar preguntas específicas y contribuir al conocimiento en el campo de la gestión ambiental.



A manera de ejemplo, se presenta a continuación una pregunta de investigación donde se integra la investigación cuantitativa y cualitativa en un caso de gestión ambiental:

Pregunta de investigación: ¿cómo afecta la actividad industrial en un área específica a la calidad del agua de un río, y cuáles son las percepciones y preocupaciones de la comunidad local al respecto?

Los resultados dentro del *enfoque cuantitativo* serían información objetiva sobre la variación en la calidad del agua en función de la actividad industrial. Con el enfoque cualitativo se realizarían entrevistas a miembros de la comunidad local para obtener una descripción detallada de las preocupaciones de la comunidad, y cómo perciben la relación entre la actividad industrial y la calidad del agua, y qué acciones son necesarias para asegurar una calidad del agua adecuada para el consumo por la comunidad.

La integración de los hallazgos para obtener una comprensión más completa y holística sería desarrollar recomendaciones basadas en la síntesis de datos cuantitativos y cualitativos para abordar preocupaciones específicas de la comunidad. Este enfoque integrado permite una comprensión más completa y contextualizada de la problemática, incorporando tanto los aspectos objetivos medidos cuantitativamente como las percepciones y preocupaciones subjetivas de la gente colectadas cualitativamente. Los resultados de esta investigación integrada permitirán no solo entender el impacto ambiental de la actividad industrial, sino el impacto social y la posibilidad de motivar a la comunidad afectada para abordar y resolver el problema.

En esta sección 1.1, hemos explorado los fundamentos y las aplicaciones de la investigación cuantitativa y cualitativa. Cada enfoque tiene sus propias fortalezas y limitaciones, y comprender cómo se complementan mutuamente es esencial para llevar a cabo significativas investigaciones.

La comprensión de los enfoques, cuantitativo y cualitativo de investigación es fundamental para desarrollar un planteamiento del problema efectivo. Al entender las características de cada enfoque, podrá determinar cuál es más apropiado para abordar su pregunta de investigación. Por ejemplo, si su interés es medir y analizar la concentración de contaminantes en un río, el enfoque cuantitativo le permitirá recopilar datos numéricos precisos y realizar análisis estadísticos. En cambio, si busca comprender cómo una comunidad percibe los impactos ambientales de una actividad industrial, el enfoque cualitativo le ayudará a explorar experiencias y perspectivas en profundidad. En algunos casos, podría ser necesario combinar ambos enfoques para obtener una comprensión más completa del problema. Esta decisión inicial sobre el enfoque de investigación guiará todo el proceso posterior, desde la formulación de objetivos hasta la selección de métodos y el análisis de resultados.

1.2. El planteamiento del problema de investigación cuantitativa

La idea de investigación es el punto de partida para cualquier proyecto de investigación. Esta idea puede surgir de diversas fuentes y procesos mentales como la observación, experiencias personales, revisión de la literatura, problemáticas del mundo real (p. ej. cambio climático), discusiones estudiante-estudiante o desafíos metodológicos. Es esencial como investigadores estar abiertos a diversas fuentes de inspiración para enfocar nuestros esfuerzos en preguntas y problemas que sean significativos y relevantes para nuestro campo de estudio.

En esta sección 1.2 comprenderemos cómo la idea de investigación evoluciona hacia un problema de investigación cuantitativa. Asimismo, revisaremos diferentes estrategias para comprender cómo esbozar los objetivos y preguntas de investigación, la justificación, la viabilidad y la evaluación de las deficiencias en el conocimiento del problema.

Plantear un problema de investigación en la gestión ambiental es esencial para entender, abordar y resolver problemas ambientales de manera sistemática y efectiva. De esta manera podemos contribuir al desarrollo sostenible y a la preservación del medioambiente.

Para estructurar de manera más formal una idea de investigación, es crucial comprender los antecedentes relevantes de la problemática ambiental o social que pretendemos resolver. Estos antecedentes proporcionan contexto, ayudan a identificar lagunas en el conocimiento existente y orientan el enfoque de la investigación. Al conocer los antecedentes, se puede establecer una base sólida para formular preguntas de investigación significativas, diseñar metodologías apropiadas y establecer la relevancia y originalidad del estudio dentro de determinado campo profesional.

El **planteamiento del problema** es fundamental para dirigir el estudio hacia objetivos claros. Este proceso implica:

- Objetivos y/o preguntas de investigación: expresan con claridad qué se pretende lograr con la investigación. Los objetivos deben ser:
 - Claros y precisos.
 - Medibles y alcanzables.
 - Congruentes entre sí.
 - Expresados con verbos en infinitivo.

Por ejemplo, un objetivo podría ser: "determinar la concentración de metales pesados en los sedimentos del río X durante el periodo 2023-2024".

- Justificación: explica por qué es necesaria la investigación, considerando:

- Valor teórico: ¿Qué vacíos de conocimiento llenará?
- Utilidad práctica: ¿Qué problemas ayudará a resolver?
- Relevancia social: ¿Quiénes se beneficiarán?
- Implicaciones ambientales: ¿Qué impacto tendrá en la conservación o gestión ambiental?



- Viabilidad: Evalúa la factibilidad de la investigación en términos de:

- Recursos disponibles.
- Tiempo requerido.
- Acceso a información y datos.
- Conocimientos y habilidades necesarias.
- Consideraciones éticas.



- Evaluación de las deficiencias en el conocimiento identifica:

- Qué se sabe actualmente sobre el tema.
- Qué aspectos no han sido estudiados.
- Qué limitaciones tienen los estudios previos.
- Qué nuevas perspectivas puede aportar la investigación.



La correcta articulación de estos elementos asegura un planteamiento sólido que guiará efectivamente el proceso de investigación (Hernández- Sampieri, 2018).

Uno de mis propósitos, por ejemplo, es encontrar las causas del abandono de perros en las ciudades de Ecuador. El abandono de perros es un problema complejo que puede tener diversas causas como la falta de compromiso por parte del propietario o problemas de comportamiento del perro (agresividad, destructividad o falta de entrenamiento). Por tanto, el abordaje de este problema requiere un enfoque integral para aplicar medidas de control y prevención.

1.3. Componentes del planteamiento de investigación cuantitativa

Piense en algo que quiera desarrollar una vez que egrese de su carrera (un negocio, un consultorio, un proyecto gubernamental, social o en una empresa, una tecnología, un producto, una labor o tarea, etc.). Planteé un problema de investigación relacionado con ello. ¿Cómo plantearía los objetivos para su negocio o empresa?



El problema de investigación se plantea a través de cinco componentes o elementos que se encuentran fuertemente relacionados entre sí.

Como revisamos anteriormente, el planteamiento del problema de investigación comprende elementos fundamentales que deben articularse de manera coherente. Recordemos brevemente estos componentes (Hernández-Sampieri, 2018):

- Los objetivos expresan con claridad lo que pretendemos lograr.
- Las preguntas de investigación que orientan hacia las respuestas que buscamos.
- La justificación que explica la relevancia y beneficios del estudio.
- La viabilidad que evalúa los recursos y condiciones necesarias.
- La evaluación de deficiencias que identifica los vacíos en el conocimiento actual.

Para comprender mejor cómo estos elementos se integran en un planteamiento efectivo, revisaremos un caso práctico en el campo de la gestión ambiental que nos permitirá visualizar cómo se aplican estos conceptos en una situación real.

Ejemplo: gestión de la calidad del agua de un río urbano

La idea de esta investigación surgió a partir de la observación directa de habitantes locales que expresaron preocupaciones sobre el estado del río que atraviesa su ciudad. Además, una revisión de informes ambientales reveló un aumento en los niveles de contaminantes. La necesidad de abordar estos

problemas y mejorar la gestión de la calidad del agua se volvió evidente, motivando así el desarrollo de esta investigación interdisciplinaria que abarca aspectos de gestión ambiental y salud pública. La colaboración con organizaciones locales de conservación del medioambiente también contribuyó a la identificación de la relevancia y la oportunidad de este proyecto. Los componentes del planteamiento de esta idea son:

1. **Objetivo de investigación:** determinar las fuentes de contaminación y las medidas de gestión existentes en la calidad del agua del río urbano.
2. **Pregunta de investigación:** ¿cómo afectan las fuentes de contaminación y las medidas de gestión existentes en la calidad del agua del río urbano?
3. **Justificación:** la calidad del agua es crucial para la salud humana y la sostenibilidad ambiental. Una comprensión profunda de los factores que afectan la calidad del agua y la evaluación de las estrategias de gestión existentes son esenciales para garantizar un suministro de agua seguro y la preservación del ecosistema fluvial.
4. **Viabilidad:** respecto a la viabilidad técnica, se cuenta con tecnologías de monitoreo avanzadas y expertos en calidad del agua. El proyecto es viable económicamente con fondos disponibles para muestreo y análisis. Sobre la viabilidad temporal, el estudio se llevará a cabo en un período de 12 meses para capturar variaciones estacionales.
5. **Evaluación de las deficiencias:** algunas posibles deficiencias en el conocimiento de la gestión de la calidad del agua de los ríos urbanos podrían incluir:
 - Falta de datos completos y actualizados sobre la calidad del agua en diferentes ríos urbanos.
 - Limitada comprensión de las fuentes y tipos específicos de contaminación que afectan a los ríos urbanos.
 - Escasez de investigación sobre los impactos a largo plazo de la contaminación del agua en la salud humana, la vida silvestre y los ecosistemas acuáticos en entornos urbanos.
 - Falta de comprensión sobre cómo los cambios climáticos y ambientales están afectando la calidad del agua en los ríos urbanos y cómo estas tendencias pueden influir en la gestión futura.



- Limitada investigación sobre tecnologías innovadoras y soluciones de ingeniería para monitorear y mejorar la calidad del agua en entornos urbanos, así como su viabilidad y eficacia en diferentes contextos.
- Falta de investigación interdisciplinaria que integre el conocimiento científico, tecnológico, social y político para abordar de manera integral los desafíos en la gestión de la calidad del agua en entornos urbanos.

El planteamiento de la idea de investigación es sólido y bien estructurado. La motivación surge de una combinación de observaciones directas, preocupaciones de la comunidad e informes ambientales, lo que destaca la relevancia y la necesidad de abordar el problema. La naturaleza interdisciplinaria del proyecto, que abarca la gestión ambiental y la salud pública, es un punto fuerte. El objetivo y la pregunta de investigación están claramente definidos y enfocados en determinar las fuentes de contaminación y evaluar las medidas de gestión existentes. La justificación resalta la importancia de la calidad del agua para la salud humana y la sostenibilidad ambiental. La viabilidad del proyecto se ha considerado desde perspectivas técnicas, económicas y temporales. La disponibilidad de tecnologías de monitoreo, expertos y fondos, junto con un cronograma de 12 meses, sugiere que el proyecto es factible. La evaluación de las deficiencias en el conocimiento actual es exhaustiva y destaca áreas clave que requieren más investigación, como datos actualizados, fuentes específicas de contaminación, impactos a largo plazo, influencias del cambio climático, soluciones innovadoras y enfoques interdisciplinarios.

Para enriquecer su comprensión sobre los objetivos de investigación, le sugiero revisar la siguiente presentación de diapositivas (PPT): [Elaborar objetivos de investigación](#).

Después de revisar la presentación, le animo a participar en el siguiente juego de arrastrar y soltar sobre la Elaboración de objetivos de investigación. Esto le permitirá fortalecer su comprensión de una manera dinámica e interactiva.

[Elaboración de objetivos de investigación](#)

Es momento de aplicar sus conocimientos a través de las actividades que se han planteado a continuación:



Actividades de aprendizaje recomendadas



1. Para comprender el proceso de elaboración de un objetivo de investigación paso a paso, le recomiendo, revisar el video "[Superfórmula para redactar objetivos generales y específicos](#)".

Como se puede observar en el video, existen objetivos específicos que permiten cumplir con el objetivo general de una investigación. En este video, además, se muestra una fórmula sencilla pero eficaz para la elaboración correcta de los objetivos. Este video es valioso para que pueda desarrollar las actividades planteadas en esta asignatura, pero sobre todo para que adquiera la habilidad de hacer los objetivos de un proyecto de investigación correctamente.

2. A continuación, le invito a desarrollar la autoevaluación para evaluar la comprensión de la primera unidad respecto a la investigación cuantitativa y cualitativa, y al planteamiento del problema de investigación.



[Autoevaluación 1](#)

Seleccione la opción correcta.

1. **Ejercicio 1:** se plantean dos casos dentro del campo de la gestión ambiental. A raíz de lo que revisó en esta unidad, identifique, ¿qué enfoque, cuantitativo, cualitativo o mixto, tiene cada caso planteado?

Caso 1: una empresa planea construir una hidroeléctrica en una zona rural. La investigación se enfocará en comprender cómo la comunidad local percibe y experimenta esta hidroeléctrica y cuáles podrían ser sus preocupaciones, expectativas y beneficios percibidos.

Los resultados esperados serían:

- Una comprensión detallada de las preocupaciones, expectativas y percepciones de la comunidad.
- Identificación de áreas clave de impacto potencial en la calidad de vida de los residentes.



Caso 2: una ciudad está experimentando cambios en la calidad del agua en su río principal debido al crecimiento urbano. Se busca evaluar los indicadores de la salud del ecosistema fluvial.

Los resultados esperados serían:

- Una evaluación precisa de la calidad del agua.
- Identificación de factores que podrían influir en las medidas de calidad del agua.
- Descripción de las percepciones de la comunidad sobre la calidad del agua, sus usos y prácticas relacionadas.

2. **Ejercicio 2:** lea y evalúe los siguientes objetivos o preguntas de investigación y responda ¿Cuál es el tipo de problema(s) o error que tienen o en qué incumplen?

Ejemplo: investigar el impacto de la deforestación en la biodiversidad del Amazonas.

El error en este objetivo de investigación es que es demasiado amplio y general. No especifica claramente qué aspectos de la biodiversidad se van a evaluar ni cómo se medirá el impacto de la deforestación.

Además, no define claramente el alcance geográfico o temporal del estudio, lo que podría dificultar la recopilación de datos y la interpretación de los resultados. Un objetivo más específico y bien definido podría ser: "Evaluar el efecto de la deforestación en la diversidad de especies de aves en una región específica del Amazonas

durante un período de cinco años, para comparar la abundancia y diversidad de aves en áreas deforestadas y áreas intactas de la selva”.



- a. Determinar los problemas de contaminación de las mecánicas.
- b. Disminuir los niveles de contaminación provocados por los automotores.
- c. ¿Cuál es el impacto ambiental en la Amazonía durante el último año de determinada empresa minera?
- d. Elaborar estrategias efectivas de conservación para especies en peligro de extinción.
- e. Aplicar un cuestionario y una prueba estadística para analizar la diferencia entre hombres y mujeres en cuanto a la percepción sobre el cambio climático.
- f. ¿De qué manera se puede educar a los niños de primaria de la ciudad de Guayaquil, Ecuador?



3. **Ejercicio 3:** imagine que ha conversado con algunos compañeros de carrera y entre todos ha surgido la siguiente idea de investigación:

“Ecuador, conocido por su diversidad biológica y ecológica, se enfrenta a desafíos ambientales en el sector agrícola debido a prácticas que podrían tener impactos en los ecosistemas locales. La idea de investigación se enfocaría en la región costera, caracterizada por la agricultura intensiva en ecosistemas sensibles. Esta investigación sería crucial para comprender el equilibrio entre la producción agrícola y la conservación ambiental en una región de alta importancia ecológica.

Además, proporcionará información valiosa para el desarrollo de estrategias sostenibles que podría ayudar tanto a los agricultores como al medioambiente de la costa de Ecuador”.

Evalúe esta idea de investigación a través de las siguientes preguntas:

- ¿El objetivo es claro, preciso y llevará a la realización de una investigación en la “realidad”?
- ¿La pregunta es ambigua?

- ¿Qué va a lograrse con este objetivo de investigación?
- ¿Es posible realizar esa investigación?
- Además, evalúe el planteamiento de su objetivo de acuerdo con los criterios expuestos en esta unidad.



Después de contestar estas preguntas, transforme la idea en un objetivo y en una pregunta de investigación cuantitativa concisa y concreta que cumpla con los criterios establecidos para elaborar los “Objetivos de la investigación cuantitativa” y “Preguntas de investigación cuantitativa”.



- a. Transforme y desarrolle la idea en un objetivo de investigación.
- b. Transforme y desarrolle la idea en una pregunta de investigación cuantitativa, concisa y concreta.



[Ir al solucionario](#)



¡Excelente trabajo! En la siguiente unidad, exploraremos más conceptos acerca de la elaboración del marco teórico.

Resultado de aprendizaje 2:

Realiza revisiones bibliográficas efectivas.

La finalidad del resultado de aprendizaje es comprender el proceso para buscar y revisar la literatura relacionada con el problema de investigación para elaborar marcos teóricos que sustenten el desarrollo de la investigación.

Contenidos, recursos y actividades de aprendizaje recomendadas

Recuerde revisar de manera paralela los contenidos con las actividades de aprendizaje recomendadas y actividades de aprendizaje evaluadas.



Semana 2

Unidad 2. El marco teórico

El *marco teórico* consiste en la revisión de teorías, conceptos y estudios previos que proporcionan la base conceptual para abordar un determinado problema de investigación. Una vez que la idea de investigación evoluciona a un objetivo de investigación, iniciamos el proceso para elaborar el *marco teórico*. Si fuéramos músicos y quisieramos componer nuevas canciones, es fundamental conocer las canciones y ritmos existentes para proponer realmente algo novedoso en la música.

2.1. Elaboración del marco teórico

La elaboración cuidadosa del marco teórico en la gestión ambiental es esencial para encajar nuestra investigación en el conjunto ("big picture") de lo que se conoce sobre un tema o tópico estudiado. No solo proporciona una estructura conceptual sólida, sino que también impulsa la investigación, orienta la toma de decisiones y contribuye al avance en la preservación y cuidado del entorno natural.

Imagine el siguiente escenario. Usted se ha graduado de la carrera de Gestión Ambiental y lo contrata un municipio local para investigar las causas principales de contaminación del principal río que atraviesa la ciudad que habita. La finalidad del alcalde es utilizar los recursos limitados de que dispone para prevenir la contaminación del río y asegurar el suministro del agua para el riego de la zona. Sin embargo, usted desconoce el historial y antecedentes de contaminación de este río. ¿Cómo puede conocer esta información para saber cuál es el primer paso que debe seguir para realizar la tarea que le han asignado? A continuación, encontrará las herramientas que le permitirán responder a esta interrogante.

El **marco teórico** sustenta toda investigación científica. Para comprender su importancia, analizaremos:

- Naturaleza del marco teórico: el marco teórico es una revisión crítica y sistemática de la literatura existente relacionada con el problema de investigación. Implica analizar y exponer teorías, enfoques, investigaciones previas y antecedentes que fundamenten el estudio (Munch Galindo, 2009).
- Función en la investigación: su desarrollo es fundamental en cualquier investigación porque:
 - Delimita el área de estudio y sus alcances.
 - Previene errores cometidos en estudios anteriores.
 - Orienta sobre cómo realizar la investigación.
 - Proporciona un marco de referencia para interpretar resultados.
- Aplicación en gestión ambiental: en el campo de la gestión ambiental, un marco teórico sólido permite:
 - Contextualizar el problema ambiental dentro del conocimiento existente.
 - Identificar variables ambientales relevantes y sus relaciones.
 - Fundamentar la interpretación de resultados.
 - Proponer soluciones basadas en la evidencia científica.



Subraye conceptos clave, definiciones y ejemplos relacionados con la construcción del marco teórico. Concluya su revisión con una breve reflexión personal sobre la importancia del marco teórico en la gestión ambiental.

2.2. Fases para la elaboración del marco teórico

Si nuestro objetivo de investigación fuera “*Evaluar la calidad del agua del río Machángara de la ciudad de Quito mediante oxígeno disuelto*”, ¿por dónde empiezo?, ¿qué proceso me puede guiar para elaborar el marco teórico de mi objetivo de investigación?

La elaboración del marco teórico es tanto un proceso sistemático como un producto final que sustenta la investigación (Hernández-Sampieri, 2018). El proceso inicia con la detección y obtención de literatura pertinente, donde el investigador debe:

- Identificar fuentes primarias relevantes.
- Consultar bases de datos académicos.
- Seleccionar términos de búsqueda apropiados.
- Evaluar la calidad de las fuentes.

La revisión analítica y selectiva de la literatura constituye una fase crucial donde se examina cuidadosamente cada fuente para extraer información relevante. Esta revisión implica evaluar la calidad y pertinencia de cada documento, establecer conexiones entre diferentes autores y perspectivas teóricas, y organizar la información de manera sistemática (Munch Galindo, 2009).

Al evaluar el panorama que revela la literatura, el investigador debe identificar (Hernández-Sampieri, 2018):

- Teorías fundamentales del campo.
- Estudios previos relevantes.
- Metodologías utilizadas.

- Vacíos en el conocimiento.

La organización y estructuración de la literatura representa una fase donde la información se ordena de manera lógica y coherente. Esta organización puede realizarse mediante mapas conceptuales o índices temáticos que guiarán la redacción posterior, asegurando una estructura que facilite la comprensión de las relaciones entre diferentes conceptos y teorías (Munch Galindo, 2009).

El proceso culmina con la redacción y revisión del marco teórico, que debe presentar una argumentación clara y fundamentada, integrando diferentes perspectivas teóricas de manera coherente. Como producto final, el marco teórico no solo sistematiza el conocimiento existente sobre el problema de investigación, sino que también establece las bases teóricas para el desarrollo del estudio y la interpretación de sus hallazgos (Hernández-Sampieri, 2018).

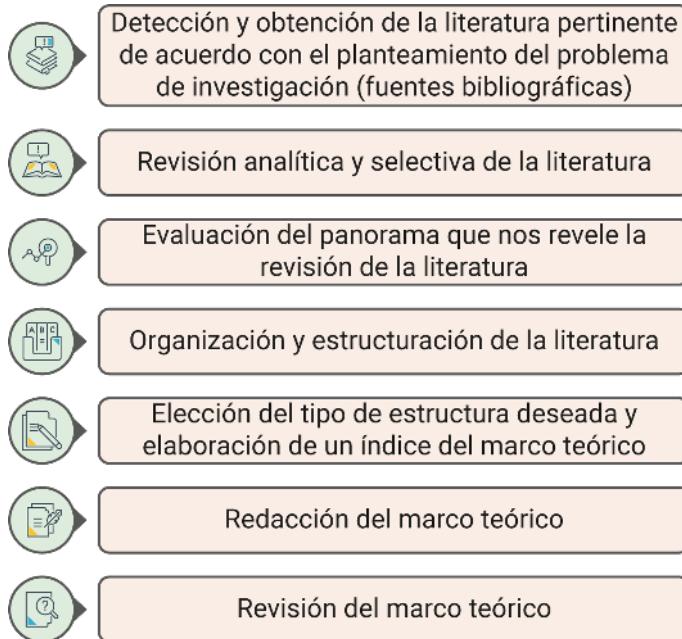
Analice el contenido de manera activa y reflexiva. Subraye conceptos y definiciones clave. Formule preguntas reflexivas sobre el contenido. Por ejemplo: ¿Cómo puedo aplicar estos conceptos a mi propia investigación en gestión ambiental?, ¿cómo el marco teórico contribuye a la comprensión de los desafíos ambientales?, ¿cuáles son las teorías más relevantes para mi área específica de investigación?

A continuación, muestra la figura 1 sobre el proceso sistemático que involucra la búsqueda, evaluación, organización y redacción del marco teórico, que es una parte fundamental de cualquier trabajo de investigación.



Figura 1

Proceso para elaborar el marco teórico (etapas).



Nota. Adaptado de *Metodología de la investigación* [Infografía], por Hernández-Sampieri, 2018, McGraw-Hill Interamericana. CC BY 4.0.

Siguiendo el sistema de etapas de la figura 1, para elaborar el marco teórico del objetivo “*Evaluar la calidad del agua del río Machángara de la ciudad de Quito mediante oxígeno disuelto*”, lo primero que deberíamos hacer sería ir a internet para detectar fuentes primarias (artículos científicos, libros, documentos, etc.) en diversos centros de referencias como Google Académico y la [base de datos UTPL](#) para identificar palabras o términos clave de búsqueda en español e inglés.

¿Qué términos debo utilizar? Con una primera búsqueda utilizando los términos de nuestro objetivo de investigación, podríamos ir afinando y actualizando los términos de búsqueda para detectar y obtener la literatura pertinente. Luego, podemos revisar las 30 primeras referencias encontradas para seleccionar cuáles nos serían útiles y cuáles no, para elaborar el marco teórico para nuestro objetivo.

La construcción del marco teórico depende de lo que encontremos en la revisión de la literatura, la cual puede revelar diferentes grados en el desarrollo del conocimiento sobre nuestro tema:

- a. Que haya una *teoría completamente desarrollada* que se aplique a nuestro objetivo de investigación.
- b. Que existan *varias teorías* que se apliquen al problema de investigación.
- c. Que haya generalizaciones o conclusiones de otros estudios que se adapten a nuestro problema.
- d. Que encontremos *descubrimientos interesantes pero parciales* que no se ajustan a una teoría.
- e. Que solamente existan guías aún no estudiadas e *ideas vagamente relacionadas* con el problema de investigación.

En cada caso, varía la estrategia para construir el marco teórico. Un marco teórico para el objetivo de investigación “*Evaluar la calidad del agua del río Machángara de la ciudad de Quito mediante oxígeno disuelto*” podría ser el siguiente:

El oxígeno disuelto es un parámetro crucial para evaluar la calidad del agua de los ríos, ya que es esencial para la vida acuática y refleja la salud general del ecosistema acuático. El oxígeno disuelto en el agua influye en numerosos procesos biológicos, incluida la respiración de organismos acuáticos, la descomposición de materia orgánica y la producción primaria (Pérez-Castillo & Rodríguez, 2008; Baque-Mite et al. 2016). Además, la cantidad de oxígeno disuelto puede variar significativamente debido a factores como la temperatura del agua, la presión atmosférica y la actividad biológica (Rojas, 2018; Muñoz, et al. 2015).

La evaluación de la calidad del agua del río Machángara en Quito, Ecuador, es crucial debido a su importancia como fuente de agua potable y su papel en el mantenimiento del equilibrio ecológico en la región (Campaña & Gualoto 2015). Según estudios previos como el de Martínez et al. (2019), el río Machángara ha experimentado problemas de contaminación debido a descargas de aguas residuales sin tratar, vertidos industriales y actividades

agrícolas cercanas (Da Ros, 1995; Terán, 2014; Campaña & Gualoto 2015). Estas fuentes de contaminación pueden afectar negativamente los niveles de oxígeno disuelto en el agua y, por lo tanto, la salud de los ecosistemas acuáticos y la seguridad del suministro de agua potable (Castro, 2014).

Por lo tanto, es fundamental llevar a cabo una evaluación detallada de la calidad del agua del río Machángara mediante la medición del oxígeno disuelto (Reinoso, et al., 2015; Campaña, et al., 2017). Esto proporcionará información crucial para la toma de decisiones en términos de gestión ambiental y protección de los recursos hídricos en la ciudad de Quito.

Referencias del marco teórico acerca de la calidad del agua del río Machángara:

- Campaña Lozano, R. A., & Gualoto Kirochka, E. (2015). *Evaluación fisicoquímica y microbiológica de la calidad del agua de los ríos Machángara y Monjas de la red hídrica del distrito metropolitano de Quito (DMQ)* (Tesis de Licenciatura).
- Campaña, A., Gualoto, E., & Chiluisa-Utreras, V. (2017). Evaluación fisicoquímica y microbiológica de la calidad del agua de los ríos Machángara y Monjas de la red hídrica del distrito metropolitano de Quito. *Revista Bionatura*, 2(2), 305-310.
- Castro Llundo, V. T. (2014). *Estudio y desarrollo de un sistema de tratamiento biológico de aguas residuales de las parroquias rurales Cusubamba y Mulalillo del cantón Salcedo de la provincia de Cotopaxi* (Tesis de Licenciatura, Universidad Técnica de Ambato. Facultad de Ciencia e Ingeniería en Alimentos. Carrera de Ingeniería Bioquímica).
- Da Ros, G. (1995). *La contaminación de aguas en Ecuador: una aproximación económica*. Editorial Abya Yala.

- Baque-Mite, R., Simba-Ochoa, L., González-Ozorio, B., Suatunce, P., Diaz-Ocampo, E., & Cadme-Arevalo, L. (2016). Calidad del agua destinada al consumo humano en un cantón de Ecuador. *Revista Ciencia UNEMI*, 9(20), 109-117.
- Muñoz, H., Orozco, S., Vera, A., Suárez, J., García, E., Neria, M., & Jiménez, J. (2015). Relación entre oxígeno disuelto, precipitación pluvial y temperatura: río Zahuapan, Tlaxcala, México. *Tecnología y ciencias del agua*, 6(5), 59-74.
- Pérez-Castillo, A. G., & Rodríguez, A. (2008). Índice fisicoquímico de la calidad de agua para el manejo de lagunas tropicales de inundación. *Revista de Biología Tropical*, 56(4), 1905-1918.
- Reinoso Chisaguano, I. C. (2015). *Evaluación ambiental del río Machángara* (Tesis de Licenciatura, Quito, 2015.).
- Rojas, H. A. (2010). Aplicación de un modelo de correlación entre las variables temperatura, densidad y oxígeno disuelto para la identificación de termoclina en la estratificación térmica. *Reportes científicos de la FACEN*, 1(2), 2-14.
- Terán, J. L. N. (2014). *Maestría en toxicología industrial y ambiental* (Tesis Doctoral, Universidad de Cuenca).

Continuemos con el aprendizaje mediante su participación en la actividad que se describe a continuación:



Actividad de aprendizaje recomendada

Estimado estudiante, para reforzar la comprensión de los contenidos de esta unidad, le invito a desarrollar la autoevaluación que a continuación se presenta.



Autoevaluación 2

1. El marco teórico consiste en la revisión de teorías, conceptos y estudios previos que proporcionan la base conceptual para abordar un determinado problema de investigación.
 - a. Verdadero.
 - b. Falso.

2. La elaboración del marco teórico no tiene influencia en el avance de la investigación ni en la toma de decisiones relacionadas con la preservación del medioambiente.
 - a. Verdadero.
 - b. Falso.

3. La detección de fuentes primarias se realiza exclusivamente a través de la base de datos UTPL.
 - a. Verdadero.
 - b. Falso.

4. La construcción del marco teórico se basa en lo que se encuentra durante la revisión de la literatura.
 - a. Verdadero.
 - b. Falso.

5. Utilizar términos específicos y relevantes en la búsqueda bibliográfica es fundamental para obtener la literatura pertinente.
 - a. Verdadero.
 - b. Falso.



6. Revisar las primeras 30 referencias encontradas durante la búsqueda bibliográfica es suficiente para seleccionar la literatura útil para la elaboración del marco teórico.
- a. Verdadero.
b. Falso.
7. La etapa 1 del proceso de elaboración del marco teórico implica la detección y obtención de la literatura pertinente según el planteamiento del problema de investigación.
- a. Verdadero.
b. Falso.
8. La etapa 5 del proceso consiste en elegir el tipo de estructura deseada y elaborar un índice del marco teórico.
- a. Verdadero.
b. Falso.
9. Respecto al objetivo de investigación "*Evaluar la calidad del agua del río Machángara de la ciudad de Quito mediante oxígeno disuelto*", busque en Google Académico, 10 referencias y extraiga la información pertinente. Luego, elija dos o más teorías que hagan referencia al mismo fenómeno y compárelas.
10. Para el objetivo "Evaluar la calidad del agua del río Machángara de la ciudad de Quito mediante oxígeno disuelto", realice una búsqueda sistemática en bases de datos académicas para identificar al menos 10 referencias científicas relevantes. A partir de estas referencias, organice la información siguiendo una estructura lógica que incluya conceptos clave, teorías fundamentales y metodologías aplicables. ¿De qué manera este proceso de búsqueda y organización de la literatura le ayudó a comprender la problemática de calidad del agua en el río Machángara y la importancia del oxígeno disuelto como indicador ambiental?



[Ir al solucionario](#)



“¡Felicitaciones por llegar hasta aquí! Ha construido un sólido marco teórico, cimiento de su conocimiento.



Resultado de aprendizaje 3:

Plantea hipótesis en torno al tema de investigación.

La finalidad del resultado de aprendizaje es comprender los factores que determinan el alcance inicial y final de un estudio, aprender a formular hipótesis y definir de manera conceptual las variables contenidas en una hipótesis.

Contenidos, recursos y actividades de aprendizaje recomendadas

Recuerde revisar de manera paralela los contenidos con las actividades de aprendizaje recomendadas y actividades de aprendizaje evaluadas.



Semana 3

Unidad 3. Los alcances en la investigación cuantitativa

Definir el alcance de la investigación, ya sea exploratorio, descriptivo, correlacional o explicativo, es una parte esencial del proceso de planificación y diseño de la investigación. El alcance de una investigación define el nivel de profundidad con que se abordará el fenómeno de estudio. La elección del alcance determinará la estrategia de investigación más apropiada para abordar problemas ambientales específicos. Los tipos de alcance según Hernández-Sampieri et., Al (2028) son:

1. Estudios exploratorios: Se realizan cuando el objetivo es examinar un tema poco estudiado o novedoso. Por ejemplo, investigar el impacto de nuevos contaminantes emergentes en ecosistemas acuáticos.
2. Estudios descriptivos: buscan especificar propiedades, características y perfiles de fenómenos ambientales. Por ejemplo, caracterizar la calidad del agua de un río mediante parámetros fisicoquímicos.
3. Estudios correlacionales: asocian variables mediante un patrón predecible. Por ejemplo, relacionar la concentración de contaminantes con la biodiversidad acuática.

4. Estudios explicativos: determinan las causas de los fenómenos estudiados.

Por ejemplo, establecer cómo los vertidos industriales afectan los niveles de oxígeno disuelto en ríos urbanos.



La elección del alcance dependerá de:

- El estado del conocimiento sobre el problema.
- La perspectiva que se pretende dar al estudio.
- Los objetivos planteados.



Las diferencias entre los distintos alcances de investigación pueden comprenderse analizando sus características distintivas y propósitos específicos:



- Alcances exploratorios vs. descriptivos: los estudios exploratorios se centran en examinar problemas poco estudiados y generar nuevas perspectivas, mientras que los descriptivos buscan detallar con precisión las características de un fenómeno ya conocido. Por ejemplo, un estudio exploratorio investigaría los efectos de microplásticos emergentes en ecosistemas acuáticos, mientras que uno descriptivo caracterizaría los niveles actuales de contaminación en un río.
- Descriptivos vs. correlacionales: los estudios descriptivos se limitan a medir y caracterizar variables de manera independiente, mientras que los correlacionales evalúan el grado de asociación entre dos o más variables. Por ejemplo, un estudio descriptivo mediría los niveles de oxígeno disuelto en un río, mientras que uno correlacional analizaría la relación entre estos niveles y la abundancia de especies acuáticas.
- Correlacionales vs. explicativos: los estudios correlacionales establecen relaciones entre variables sin determinar causalidad, mientras que los explicativos buscan establecer las causas de los fenómenos. Por ejemplo, un estudio correlacional podría encontrar una relación entre la temperatura del agua y el oxígeno disuelto, mientras que uno explicativo determinaría cómo y por qué la temperatura afecta los niveles de oxígeno.



Existe una progresión natural entre los alcances:

- Los exploratorios preparan el terreno para nuevos estudios.
- Los descriptivos fundamentan las investigaciones correlacionales.
- Los correlacionales proporcionan información para estudios explicativos.
- Los explicativos son la base para la comprensión profunda de los fenómenos.



Revise las definiciones y características de los alcances exploratorio, descriptivo, correlacional y explicativo. Observe cómo se selecciona un alcance específico según los objetivos del estudio. Identifica las diferencias entre los alcances. Pregúntese cómo se distingue un estudio exploratorio de uno descriptivo, y así sucesivamente. Resalta las particularidades que caracterizan a cada tipo de investigación.



Pregúntese: ¿cuándo sería más apropiado elegir un alcance exploratorio en una investigación?, ¿cómo se relaciona la naturaleza de la pregunta de investigación con la elección del alcance?, ¿cuáles son las limitaciones de cada tipo de alcance?

A continuación, se ofrecen ejemplos de cada alcance de investigación relacionados con la gestión ambiental.

3.1. Alcance exploratorio

Los estudios con alcance exploratorios tienen como propósito familiarizarnos con un planteamiento desconocido, poco estudiado o novedoso.

Un ejemplo sería, explorar los factores que pueden estar impactando la calidad del agua del río de nuestra ciudad, identificando posibles fuentes de contaminación y fenómenos relevantes. Este estudio exploratorio proporciona una visión integral de los factores impactantes en la calidad del agua del río, generando hipótesis iniciales que puedan orientar estudios posteriores más

detallados con alcance descriptivo, correlacional o explicativo. Aunque igualmente se podrían medir algunos parámetros de forma preliminar, como estudio piloto, para tener una idea de si la calidad del agua es adecuada o no.

3.2. Alcance descriptivo

Los estudios con alcance descriptivo sirven para analizar cómo es y cómo se manifiesta un fenómeno o planteamiento y sus variables.

Un ejemplo de estudio descriptivo sería describir la calidad del agua en el tramo urbano del río mencionado en el ejemplo anterior, identificando los principales parámetros físicos, químicos y biológicos presentes en el agua. Este estudio se enfocará exclusivamente en describir, por ejemplo, la calidad del agua en el tramo urbano del río, sin abordar relaciones de causalidad o explicaciones detalladas de los fenómenos observados. En este tipo de estudio con alcance descriptivo, los resultados podrían ser la caracterización de los niveles de contaminantes químicos, la concentración de microorganismos indicadores, los niveles de oxígeno disuelto, o los patrones estacionales en la calidad del agua del río.

3.3. Alcance correlacional

Los estudios con alcance correlacional tienen como objetivo investigar la relación o asociación entre diferentes hechos, fenómenos, conceptos, variables o características, o incluso determinar si no existe relación entre ellos.

Por ejemplo, retomando el caso anterior, podríamos realizar un estudio con alcance correlacional para analizar la relación entre la calidad del agua del río mediante oxígeno disuelto (mg/L) y la diversidad de macroinvertebrados (como chinches, escarabajos, crustáceos, moluscos, sanguijuelas y planarias). La posible correlación sería que, a mayor oxígeno disuelto, mayor diversidad de macroinvertebrados en el agua del río (Croijmans et al., 2021). Estos



resultados podrían proporcionar indicadores valiosos para la gestión ambiental y la toma de decisiones relacionadas con la calidad del agua en entornos urbanos.

3.4. Alcance explicativo

Los estudios explicativos se enfocan en descubrir las causas que generan ciertos fenómenos o establecer relaciones de causalidad entre variables. Es importante comprender la diferencia entre estudios con alcance correlacional y explicativo. En un estudio con alcance explicativo, el objetivo principal es ir más allá de simplemente identificar la relación entre variables para explicar por qué existe esa relación. El objetivo es *buscar establecer relaciones causales y comprender los mecanismos ocultos*.

Por ejemplo, un estudio correlacional podría analizar la asociación entre la cantidad de basura arrojada en un río y los niveles de contaminación del agua. Podría revelar una correlación estadística, pero no podría explicar por qué la basura y la contaminación están relacionadas. En cambio, un estudio explicativo que examina cómo la cantidad de residuos industriales vertidos en un río (variable independiente) afecta directamente los niveles de oxígeno disuelto en el agua (variable dependiente). También se podría examinar esta relación bajo condiciones de laboratorio donde se controlan todos los factores aparte de la variable independiente de interés. Así, se busca explicar cómo y por qué los residuos industriales causan cambios en los niveles de oxígeno disuelto.

Continuemos con el aprendizaje mediante su participación en las actividades que se describen a continuación:



Actividades de aprendizaje recomendadas

1. Con el fin de comprender los principales alcances de investigación le invito a revisar el video "[Ejemplos de alcances de una investigación](#)".



Como se puede observar, la investigación exploratoria tiene como objetivo principal familiarizarse con un tema poco conocido, generando ideas y formulando hipótesis. La investigación descriptiva busca caracterizar fenómenos y detallar sus características esenciales. La investigación correlacional busca establecer relaciones y asociaciones entre variables, sin intervenir en el entorno. La investigación explicativa tiene como meta comprender las causas y efectos de fenómenos, buscando explicaciones causales y evaluando variables predictoras. Cada enfoque cumple un papel único en el proceso investigativo, desde la familiarización inicial hasta la identificación y comprensión de relaciones y causas.

2. Estimado estudiante, para afianzar la comprensión de los contenidos de esta unidad, le invito a desarrollar la autoevaluación que a continuación se presenta.



Autoevaluación 3

Responda verdadero o falso según corresponda.

1. Establecer el alcance nos permite una guía clara sobre la metodología a utilizar.
 - a. Verdadero.
 - b. Falso.
2. Los estudios exploratorios tienen como objetivo familiarizarnos con planteamientos desconocidos, poco estudiados o novedosos.
 - a. Verdadero.
 - b. Falso.
3. Los estudios con alcance descriptivo buscan comprender el porqué y el cómo de un fenómeno o planteamiento.
 - a. Verdadero.

- b. Falso.
4. Los estudios con alcance correlacional buscan establecer relaciones causales entre variables.
- a. Verdadero.
- b. Falso.
5. En un estudio con alcance explicativo, se busca comprender los mecanismos ocultos detrás de las relaciones entre variables.
- a. Verdadero.
- b. Falso.
6. Ejercicio 1: a continuación, lo incito a desarrollar el siguiente ejercicio respecto al objetivo de investigación que planteamos en la unidad 3, *"Evaluar la calidad del agua del río Machángara de la ciudad de Quito mediante oxígeno disuelto"*, ¿a qué alcance inicial corresponde? Explique su respuesta. Discuta las explicaciones con su profesor o compañeros de clase.
7. Ejercicio 2: identifique el alcance de investigación al que corresponden las siguientes preguntas de investigación.
- a. ¿Cuáles son las características principales de la calidad del aire en la zona industrial de Parque Industrial El Sauce, Guayaquil, durante los meses de noviembre y diciembre en términos de niveles de partículas suspendidas?
- b. ¿Cuál es el efecto de la densidad vehicular en la concentración de partículas suspendidas en el aire en áreas urbanas de la ciudad de Cuenca, durante diferentes horas del día?
- c. ¿Qué percepciones tienen los jóvenes ecuatorianos nacidos en la provincia de Manabí sobre el cambio climático?, ¿cómo consideran que podrá impactar sus vidas en los próximos años?



d. ¿Cuáles son las razones por las que la gente abandona a los perros aumentando la población de perros callejeros en las ciudades de Ecuador?

[Ir al solucionario](#)



Felicitaciones por completar el estudio acerca de los alcances de investigación. Ahora, prepárese para dar un paso adelante hacia la formulación de hipótesis.

Contenidos, recursos y actividades de aprendizaje recomendadas



Semana 4

Unidad 4. Formulación de hipótesis y diseños de investigación

4.1. Formulación de hipótesis de acuerdo con el alcance del estudio

La formulación de hipótesis sirve de guía para todas las fases de una investigación, desde la elaboración del marco teórico hasta el análisis e interpretación de resultados. En gestión ambiental, las hipótesis pueden servir para orientar intervenciones, normativas y políticas. Una hipótesis sobre cómo ciertas prácticas agrícolas afecta la calidad del suelo podría guiar la implementación de prácticas sostenibles.

Por ejemplo, en un estudio realizado en el noreste de China, la hipótesis que guió la investigación fue: "Con una práctica agrícola integrada en un sistema de cultivo de maíz, cambiando el espaciado entre hileras, con labranza cero y retorno de residuos, aumentará la materia orgánica del suelo, la biomasa microbiana y las actividades enzimáticas en comparación con las prácticas convencionales". Luego de 12 años de estudio, los resultados mostraron que la materia orgánica del suelo con una práctica agrícola integrada aumentó. Además, la práctica agrícola integrada mejoró la biomasa microbiana debido a

un aumento en la comunidad microbiana del suelo, determinando que una práctica agrícola integrada es una tecnología de gestión viable para la agricultura sostenible (Sun, et al., 2018).

Pero, ¿qué es *la formulación de hipótesis*? Las hipótesis son explicaciones tentativas sobre el problema de investigación o fenómeno estudiado que se formulan como proposiciones o afirmaciones. La hipótesis se formula antes de llevar a cabo la investigación y sirve como guía para la recopilación y análisis de datos. Por ejemplo, si la pregunta de investigación es ¿La cantidad de luz afecta el crecimiento de las plantas? La hipótesis sería: “*Existe una relación positiva entre la cantidad de luz y el crecimiento de las plantas*”. Esta declaración establece una suposición o una posible relación entre dos variables.

Las **hipótesis** son explicaciones tentativas del fenómeno que se investiga, formuladas como proposiciones. Son guías precisas hacia el problema de investigación que indican lo que se está buscando o tratando de probar (Hernández-Sampieri, 2018; Munch Galindo, 2009).

- Necesidad de hipótesis en investigaciones cuantitativas: no todas las investigaciones cuantitativas requieren hipótesis. Su formulación depende del alcance inicial del estudio:
 - Los estudios descriptivos solo llevan hipótesis cuando predicen un valor o dato.
 - Los estudios correlacionales y explicativos siempre requieren hipótesis.
 - Los estudios exploratorios generalmente no formulan hipótesis.
- Variables en las hipótesis: las variables son propiedades que pueden fluctuar y cuya variación es susceptible de medirse u observarse. Pueden ser:
 - Independientes: causa o variable explicativa.
 - Dependientes: efecto o variable de respuesta.

- Características de las hipótesis. Deben:

- Referirse a una situación real.
- Ser comprensibles, precisas y concretas.
- Establecer relaciones entre variables.
- Ser susceptibles de verificación.
- Estar vinculadas con técnicas disponibles para probarlas.



- Tipos de hipótesis de investigación:



- a. Hipótesis predictivas: predicen un valor o dato en un tiempo específico.

Ejemplo. "La temperatura promedio del río X aumentará 2 °C para el año 2025".

- b. Hipótesis correlacionales: especifican relaciones entre dos o más variables. Ejemplo. "A mayor temperatura del agua, menor concentración de oxígeno disuelto".

- c. Hipótesis causales. Establecen relaciones de causa-efecto. Ejemplo. "La deforestación provoca un aumento en la erosión del suelo".

La importancia de las hipótesis radica en que guían el estudio, proporcionan explicaciones tentativas y apoyan la prueba de teorías. En gestión ambiental, permiten proponer soluciones tentativas a problemas ambientales que luego pueden verificarse mediante la investigación científica (Hernández-Sampieri, 2018).

Revise las definiciones y características fundamentales de una hipótesis de investigación. Preste atención al propósito de una hipótesis y su relación con los objetivos de investigación.



Formule preguntas reflexivas sobre la formulación de hipótesis como, por ejemplo: ¿cuáles son los elementos esenciales que debe tener una hipótesis?, ¿cómo se relaciona la hipótesis con la evidencia empírica y la toma de decisiones en la investigación?

A continuación, se muestra un ejemplo de hipótesis correlacional en el campo de la gestión ambiental.

Pregunta de investigación: ¿la disminución de oxígeno disuelto (mg/L) reduce la diversidad de macroinvertebrados del río Machángara de la ciudad de Quito?

1. **Identificación de variables:** variable independiente es el oxígeno disuelto y la variable de respuesta o dependiente es la diversidad de macroinvertebrados.
2. **Establezca la relación y elabore la hipótesis:** a menor cantidad de oxígeno disuelto (mg/L) menor diversidad de macroinvertebrados en el río Machángara de la ciudad de Quito. Esta es una hipótesis correlacional del tipo "a menor X, menor Y".
3. Mediante la medición del oxígeno disuelto y captura de macroinvertebrados del río Machángara se podrá verificar lo establecido en la hipótesis.

4.2. Tipos de diseño en la investigación cuantitativa

Una vez que hemos concretado el objetivo o pregunta de investigación, el marco teórico, el alcance de la investigación y la hipótesis, el siguiente paso es construir un plan de investigación para obtener la información que necesitamos para responder al planteamiento del problema.

En la ruta cuantitativa, el plan de investigación se utiliza para analizar la certeza de las hipótesis o responder a las preguntas de investigación exploratorias o descriptivas. Los planes o diseños pueden ser experimentales y no experimentales.

Imaginemos que, como gestores ambientales, estamos encargados del proyecto "Calidad del agua del río Machángara" en donde la hipótesis de investigación es "A menor cantidad de oxígeno disuelto (mg/L) menor diversidad de macroinvertebrados en el río Machángara de la ciudad de Quito". ¿Qué diseño debemos aplicar?, ¿será un diseño experimental o un diseño no experimental? A continuación, revisaremos algunas herramientas para poder responder estas interrogantes.



El **diseño de investigación** es el plan o estrategia que se desarrolla para obtener la información requerida y responder al planteamiento del problema. En la ruta cuantitativa, los diseños pueden ser experimentales o no experimentales (Hernández-Sampieri et al. 2018).

- Características y requisitos de los **diseños experimentales**:
 - a. Manipulación de variables independientes: la manipulación puede darse de diferentes formas:
 - Presencia-ausencia: por ejemplo, comparar un grupo de plantas expuestas a un contaminante versus un grupo control sin exposición.
 - Más de dos grados: cómo evaluar diferentes concentraciones de un contaminante (baja, media, alta) en el crecimiento de plantas.
 - Los retos en la manipulación incluyen:
 - Control de condiciones ambientales.
 - Estandarización de procedimientos.
 - Mantenimiento de condiciones constantes.
 - b. Medición de variables dependientes. Es crucial definir (Melo et al., 2007):
 - Qué se va a medir (por ejemplo, niveles de oxígeno disuelto).
 - Cómo se medirá (instrumentos y métodos).
 - Cuándo se realizarán las mediciones.
 - c. Control y validez interna. Para lograr control experimental se requiere:
 - Grupos equivalentes.
 - Ambiente experimental controlado.
 - Control de variables extrañas.
 - Asignación aleatoria de unidades experimentales.

Por ejemplo, en un estudio sobre el efecto de contaminantes en plantas, todos los especímenes deben recibir las mismas condiciones de luz, agua y nutrientes, variando únicamente el nivel del contaminante estudiado.

Por otro lado, la **investigación no experimental** en la ruta cuantitativa estudia fenómenos en su contexto natural, sin manipulación deliberada de variables. El investigador observa y analiza los fenómenos tal como ocurren en su ambiente (Munch Galindo, 2009). Estos diseños son particularmente útiles para estudiar fenómenos que no pueden ser manipulados por razones éticas o prácticas. En la investigación no experimental existen dos tipos de diseños, la investigación transversal y la longitudinal (Hernández-Sampiero et al., 2018).

1. La *investigación transeccional o transversal*, recolecta datos en un momento único, como una fotografía del fenómeno. Por ejemplo:

- Evaluar la calidad del aire en diferentes zonas urbanas en un día específico.
- Medir la concentración de contaminantes en varios puntos de un río simultáneamente.
- Analizar la biodiversidad en diferentes áreas de un bosque en una temporada determinada.

2. La *investigación longitudinal o evolutiva*, analiza cambios a través del tiempo en determinadas variables o en las relaciones entre ellas. Por ejemplo:

- Monitorear cambios en la calidad del agua de un río durante varios años.
- Estudiar la evolución de especies indicadoras en un ecosistema a lo largo del tiempo.
- Dar seguimiento a los efectos del cambio climático en un ecosistema específico.

Los diseños longitudinales pueden ser de:

- Tendencia: analizan cambios en una población general. Por ejemplo, estudiar cómo ha cambiado la calidad del aire en una ciudad durante las últimas dos décadas, aunque no sean exactamente los mismos puntos de medición.
- Evolución del grupo: estudian cambios en subgrupos específicos. Por ejemplo, monitorear cómo cambia la resistencia a contaminantes en

diferentes grupos de especies acuáticas (peces, anfibios, invertebrados) en un río a lo largo de cinco años.

- Panel: siguen a los mismos casos o participantes a lo largo del tiempo. Por ejemplo, monitorear exactamente los mismos árboles de una especie forestal para evaluar su respuesta al cambio climático durante una década, registrando cambios en su crecimiento, fenología y estado fitosanitario.

La elección entre un diseño transeccional o longitudinal dependerá de:

- Los objetivos de la investigación.
- El tiempo disponible para el estudio.
- Los recursos con que se cuenta.
- La naturaleza del fenómeno estudiado.

Para revisar estos contenidos le recomiendo las siguientes estrategias de estudio:

- Anote las características esenciales que definen un diseño de investigación.
- Analice los requisitos y características distintivas de los diseños experimentales.
- Examine los tipos de diseños no experimentales, enfocándose en la investigación transeccional y la investigación longitudinal.
- Formúlese preguntas reflexivas como: ¿cómo diferenciaría los diseños experimentales de los no experimentales?, ¿cuáles son los retos asociados con la manipulación de variables independientes en un experimento?, ¿por qué es esencial mantener la validez interna en un diseño experimental?, ¿en qué situaciones preferirías utilizar un diseño transeccional en lugar de un diseño longitudinal?

4.2.1. Diferencias entre un diseño experimental y no experimental

Mediante un ejemplo, vamos a comprender las principales diferencias entre un diseño experimental y uno no experimental.

El siguiente estudio tiene un **enfoque experimental** y su objetivo es evaluar el efecto de la rotación de cultivos en la productividad agrícola y la sostenibilidad ambiental en una finca agrícola durante un período de dos años. Para este estudio se selecciona una finca agrícola con condiciones similares de suelo y clima. Se establecerán dos grupos de parcelas: un grupo donde se aplicará la rotación de cultivos y otro grupo donde se mantendrá un monocultivo. En el grupo de rotación de cultivos, se establece un ciclo de rotación que incluya cultivos variados como maíz, frijol, y quinua. En el grupo de monocultivo, se siembra el mismo cultivo en cada ciclo. Finalmente, se realiza un análisis comparativo de los datos recolectados entre los dos grupos para determinar el impacto de la rotación de cultivos en la productividad y sostenibilidad agrícola.

El mismo objetivo podría abordarse mediante un diseño no experimental observacional. En este caso, el investigador recopilaría datos de diferentes fincas agrícolas que ya están aplicando prácticas de rotación de cultivos o monocultivo, sin intervenir en las decisiones de los agricultores. Se medirían variables como la productividad de los cultivos, la salud del suelo y la biodiversidad de las fincas en ambos contextos (rotación y monocultivo). Posteriormente, se realizaría un análisis correlacional para identificar patrones y relaciones entre las prácticas agrícolas observadas y los indicadores de sostenibilidad.

Es hora de reforzar los conocimientos adquiridos resolviendo las siguientes actividades:



Actividades de aprendizaje recomendadas

1. A modo de resumen, le recomiendo revisar el video acerca de "[Diseño experimental y no experimental](#)".

Como se puede revisar, la distinción clave entre diseño experimental y no experimental radica en la manipulación de variables. En el diseño experimental, se implementa una intervención controlada para

observar su efecto sobre las variables dependientes, permitiendo inferencias causales. En contraste, en el diseño no experimental, no hay manipulación deliberada; se observan las variables tal como se presentan naturalmente. Aunque el diseño no experimental, es más observacional, ambos enfoques son valiosos, ya que el primero busca establecer relaciones causales, mientras que el segundo se centra en la observación y descripción de fenómenos sin intervención directa. La elección entre ellos depende de los objetivos de investigación y la naturaleza del fenómeno estudiado.

2. Estimado estudiante, le invito a realizar la autoevaluación de esta unidad de estudio donde se presentan algunos ejercicios que le pueden ayudar para su estudio y comprensión de los contenidos. Esta actividad representa una oportunidad para reflexionar sobre su aprendizaje e identificar áreas de mejora.



Autoevaluación 4

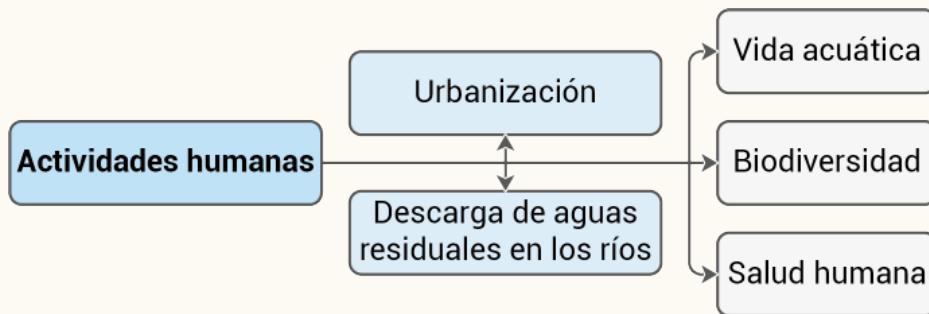
1. Ejercicio 1: arrastre cada hipótesis de investigación en su tipo correspondiente de hipótesis.

- De comparación de grupos.
- Correlacional.
- Causal.
- Correlacional.
- De pronóstico de un dato.

Hipótesis	Tipo de hipótesis
"El aumento de la deposición de nitrógeno se encuentra asociada con la disminución de la diversidad vegetal en los ecosistemas secos".	
"El porcentaje de insectos polinizadores a nivel mundial se reducirá en un 61 % en los próximos 20 años".	

Hipótesis	Tipo de hipótesis
"El aumento de la temperatura combinado con una reducción del <i>habitat</i> de las plantas con flores reducirá la cantidad de insectos a nivel mundial".	
"El barrio urbano A, en donde se implementó un programa integral de gestión de residuos sólidos, presentará una reducción en la cantidad de residuos no reciclables que el barrio B donde no se implementó dicho programa".	
"La disminución en la calidad del agua aumenta cada vez más los casos de diarrea y desnutrición crónica en la primera infancia en zonas rurales de la sierra ecuatoriana".	

2. Ejercicio 2: el siguiente esquema ilustra algunas variables contempladas en el proyecto "Calidad del Agua del Río Machángara". Usando los criterios para la redacción de una hipótesis de investigación explicados en esta sección, formule dos hipótesis concretas que se podrían proponer con base en este esquema:



Nota. Vélez, D., 2024.

3. Ejercicio 3: **Un determinado** proyecto tiene como objetivo evaluar el impacto de las estrategias de sensibilización ambiental en la gestión de residuos sólidos urbanos en la ciudad de Guayaquil. Para cumplir con este objetivo se seleccionarán 20 hogares en dos zonas urbanas

no contiguas de la ciudad de Guayaquil con las mismas características socioeconómicas. En una zona escogida aleatoriamente se implementará un programa de sensibilización ambiental y en la otra zona no se implementará ningún programa de sensibilización (grupo control). El programa de sensibilización proporcionará información mediante talleres y campañas educativas sobre la gestión adecuada de residuos. Se medirá los cambios en comportamientos de disposición y actitudes antes y después de la intervención en las dos zonas urbanas. Se utilizarán encuestas y datos de recolección de residuos para evaluar los resultados.



Analice y responda:

- a. ¿Cuál es la pregunta de investigación de este proyecto?
- b. ¿Cuál es la hipótesis que se busca probar por medio de los resultados del experimento?
- c. ¿Cuál es la variable dependiente y cuál es la variable independiente?
- d. ¿Cuántos grupos se incluyen en el experimento?
- e. ¿Son equivalentes?
- f. ¿Es un diseño experimental o no experimental?
- g. ¿Cree que este experimento pueda tener algún efecto importante?

4. Ejercicio 4: Con base en el ejercicio 3 (ejercicio anterior), ¿cómo podría plantear el mismo estudio, pero con un diseño no experimental?

[Ir al solucionario](#)



¡Felicitaciones! Hemos caminado por complicados senderos de diseños de investigación. Prepárese para comprender acerca de la importancia de la estadística. ¡Nos vemos en la siguiente unidad!

Resultado de aprendizaje 4:

Levanta, analiza e interpreta información de campo.

La finalidad del resultado de aprendizaje es aprender cómo clasificar variables, cómo usar gráficas y medidas numéricas para describir y construir una imagen mental de los datos. Además, el estudiante aprenderá las técnicas que se pueden emplear al describir estas variables para identificar posibles relaciones entre variables.

Contenidos, recursos y actividades de aprendizaje recomendadas

Recuerde revisar de manera paralela los contenidos con las actividades de aprendizaje recomendadas y actividades de aprendizaje evaluadas.



Semana 5

Después de haber establecido el marco teórico, definido los objetivos, planteado las hipótesis y delineado el diseño de nuestra investigación en gestión ambiental, es crucial pasar a la fase de recolección y análisis de datos. Aquí es donde la estadística descriptiva juega un papel fundamental en nuestra investigación.

La estadística descriptiva nos permite explorar y comprender los datos que recolectamos durante nuestro estudio ambiental. Nos brinda las herramientas necesarias para resumir, organizar y presentar estos datos de manera significativa. A través de técnicas como el cálculo de medidas de tendencia central, como la media y la mediana, y medidas de dispersión, como la desviación estándar, podemos obtener una visión clara de la distribución y variabilidad de nuestras variables de interés.

Además, la estadística descriptiva nos ayuda a identificar patrones, tendencias y relaciones preliminares en nuestros datos ambientales. Al utilizar gráficos y tablas descriptivas, podemos visualizar la información de manera efectiva, lo que facilita la interpretación y la comunicación de nuestros hallazgos.

Unidad 5. Introducción a la estadística descriptiva y tipos de variables



A partir de esta unidad exploraremos los fundamentos de la estadística descriptiva, incluyendo conceptos clave como la población y muestra, los tipos de variables y la representación de datos.

5.1. Introducción a la estadística descriptiva

En un mundo donde cada vez se genera mucha información, la estadística descriptiva puede ser la brújula que nos guía a través de un mar de datos. Imagine por un momento que usted es un investigador social tratando de entender las preferencias de música en una comunidad universitaria. Aquí la estadística se convierte en su aliada.

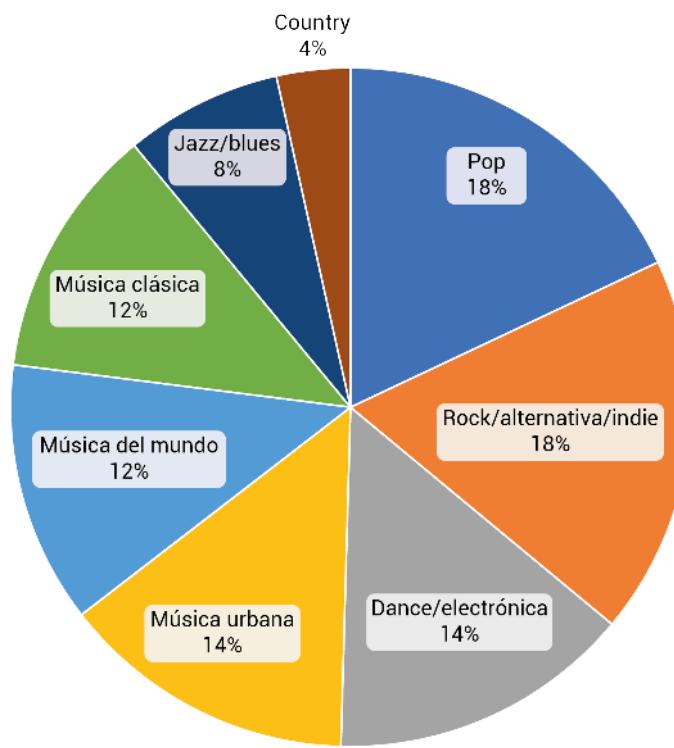
Supongamos que hemos encuestado a 200 estudiantes y les hemos preguntado acerca de sus géneros musicales favoritos. Ahora, estas respuestas son como las piezas de un rompecabezas, pero ¿cómo las organizamos para comprender el panorama completo? Aquí entra en juego la *estadística descriptiva*. La *estadística descriptiva* proporciona herramientas y técnicas para explorar, resumir y comunicar información sobre los datos observados, lo que facilita su comprensión y análisis. En lugar de enfrentarnos a esta cantidad de respuestas individuales, podemos utilizar medidas resumidas para dar sentido a la diversidad de preferencias musicales. La media o el promedio, utilizado alguna vez sin lugar a duda, nos ofrece un punto central, indicando el género musical que, *en promedio*, predomina en la comunidad.

Pero, ¿qué hay de la diversidad? Aquí la *desviación estándar* se convierte en nuestra herramienta. La desviación estándar es una medida de dispersión que indica cuánto se desvían, en promedio, los valores de un conjunto de datos respecto a la media. En nuestro ejemplo, nos muestra cuánto se alejan las preferencias individuales del promedio, revelando la amplitud del espectro musical en la comunidad.

Aunque la estadística descriptiva no es solo números, es también el arte de contar historias visuales. Por ejemplo, los *histogramas* pueden pintar un cuadro real de la distribución de preferencias, mientras que las *gráficas de pastel* nos permiten visualizar la proporción o porcentaje de cada género musical de manera intuitiva (ver figura 2).

Figura 2

Gráfica de pastel con porcentajes de las preferencias de géneros musicales por parte de estudiantes universitarios.



Nota. Adaptado de *Los géneros musicales que el mundo está escuchando* [Ilustración], por Mena, M., 2021, [statista](#), CC BY 4.0.

Ahora, ¿por qué deberíamos preocuparnos por esto? La estadística no solo es útil para investigadores sociales, sino que forma una herramienta imprescindible para el análisis de datos en prácticamente todas las

disciplinas. Los científicos la emplean para resumir resultados de experimentos, los economistas para entender tendencias de mercado, y los profesionales de la salud para analizar datos clínicos.

Así que, ya sea que estemos explorando preferencias musicales o desentrañando los principales factores de cambio global, la estadística es una herramienta que nos permite transformar datos aparentemente caóticos en conocimiento valioso. Prepárese para adentrarnos en el fascinante mundo de los datos, donde cada número cuenta una historia y la estadística es la herramienta que nos ayuda a narrarla.

La **estadística** es la ciencia que proporciona un conjunto de métodos para recolectar, organizar, analizar e interpretar datos, facilitando la toma de decisiones en situaciones de incertidumbre. La estadística es fundamental para comprender fenómenos ambientales complejos y tomar decisiones basadas en evidencia (Mendenhall et al., 2015; Melo et al., 2007).

La **población** comprende todos los elementos que comparten características específicas de interés para el estudio. Sin embargo, estudiar a toda la población suele ser impráctico o imposible. Por ello, se selecciona una **muestra**, que es un subconjunto representativo de la población. Por ejemplo, para estudiar la contaminación de un río, la población serían todas las posibles muestras de agua del río, pero se selecciona una muestra de puntos específicos para su análisis (Douglas, 2003; Mendenhall et al., 2015).

La **estadística descriptiva** se enfoca en organizar, resumir y presentar datos de manera significativa usando tablas, gráficos y medidas numéricas. Por ejemplo, calcular la concentración promedio de contaminantes en un río y representar su distribución mediante gráficos. La **estadística inferencial** utiliza datos muestrales para hacer predicciones o inferencias sobre una población más grande. Por ejemplo, usar datos de una muestra de árboles para estimar la salud general de todo un bosque (Gutiérrez Pulido & Vara Salazar, 2012; Mendenhall et al., 2015).

Por otro lado, la **descripción de datos mediante gráficas** es una herramienta visual que permite identificar patrones, tendencias y relaciones en los datos. Las principales incluyen:

- Gráficos de barras: para comparar categorías.
- Histogramas: para mostrar distribución de datos continuos.
- Gráficos de líneas: para mostrar tendencias temporales.
- Gráficos de dispersión: para explorar relaciones entre variables.

Por ejemplo, un gráfico de líneas podría mostrar cómo cambia la concentración de oxígeno disuelto en un río a lo largo del tiempo.

Las **variables** son características que pueden medirse u observarse y que pueden tomar diferentes valores. Los **datos** son los valores específicos que toman estas variables. En estudios ambientales, las variables pueden incluir temperatura, pH, concentración de contaminantes, biodiversidad, etc. (Gavilánez Luna, 2021).

Existen dos tipos de variables, las variables cualitativas y cuantitativas. Luego se subdividen cada una de ellas.

- Variables cualitativas (categóricas):

- Nominales: como tipos de ecosistemas.
- Ordinales: como niveles de contaminación (bajo, medio, alto).

- Variables Cuantitativas:

- Discretas: como número de especies en un área.
- Continuas: como temperatura o pH.

Por ejemplo, en un estudio de calidad del agua, podríamos tener:

- Variables cualitativas: tipo de cuerpo de agua (río, lago, estero).
- Variables cuantitativas: temperatura (°C), pH, oxígeno disuelto (mg/L).

Siga estas pautas para revisar estos temas y fortalecer sus habilidades estadísticas:

- Analice la definición de la estadística y reflexione sobre su importancia en la toma de decisiones en el campo de la gestión ambiental.
- Anote ejemplos prácticos de situaciones donde la estadística desempeña un papel crucial.
- Hágase preguntas reflexivas como:
 - ¿Cómo aplicaría la estadística a problemas de nuestro interés dentro de la gestión ambiental?
 - ¿Cuál es la importancia de seleccionar una muestra representativa para analizar datos?
 - ¿Cómo elegiría la gráfica más adecuada para representar datos de diferentes tipos de variables en un informe?

A continuación, se muestran algunos ejemplos de los conceptos revisados en esta unidad relacionados con la gestión ambiental.

5.2. Población y muestra

Supongamos que queremos estudiar la altura de todos los estudiantes de Gestión Ambiental de la UTPL (población). Analizar la altura de cada estudiante puede ser un poco complicado, así que seleccionamos al azar a 100 estudiantes, los cuales constituyen una muestra, y medimos sus alturas. Los resultados obtenidos de esta muestra se utilizan para hacer deducciones o inferencias sobre la altura promedio de todos los estudiantes de Gestión Ambiental de la UTPL.

En la estadística, la *población* es el grupo completo que se quiere estudiar, mientras que la *muestra* es un subconjunto representativo de esa población. La estadística utiliza la información de la muestra para hacer afirmaciones más amplias sobre la población en su conjunto.

5.3. Variables y datos

En **estadística**, una *variable* es cualquier característica, cantidad o propiedad que puede medirse o clasificarse y que puede variar entre diferentes individuos, elementos o eventos en un estudio. Es decir, es algo que puede cambiar y que estamos interesados en medir. Por ejemplo, una variable puede ser el género de una persona (masculino o femenino), el número de horas de estudio, o la calificación obtenida en una determinada asignatura.

Como parte del diseño de investigación, es fundamental definir previamente cuál será la *unidad experimental* del estudio o experimento. La *unidad experimental* es el elemento sobre el cual se realiza la observación, medición o manipulación. Puede ser una persona, un gen, una parcela del campo, un ensayo, un río, un paisaje o incluso un conjunto de paisajes.

Ejemplo de unidad experimental en gestión ambiental:



Imagine que está llevando a cabo un estudio en una comunidad urbana para evaluar el impacto de la implementación de un programa de reciclaje en la reducción de residuos sólidos. La *unidad experimental* en este caso sería *cada hogar individual* dentro de una comunidad que participa en el programa de reciclaje.

5.4. Tipos de variables

La clasificación de variables en diferentes tipos tiene que ver con la naturaleza y el nivel de medición de la información que representan. La distinción entre los tipos de variables es esencial porque determina el tipo de análisis estadístico apropiado y las operaciones matemáticas que pueden aplicarse a los datos. Las variables se clasifican en *cualitativas* y *cuantitativas*. Un ejemplo de variable cualitativa en la gestión ambiental puede ser la “*calidad del aire*” y sus categorías serían: bueno, moderado, y malo. Un ejemplo de variable cuantitativa sería, en cambio, la cantidad de *dióxido de carbono* (CO_2) emitido por una planta industrial en toneladas por año.



Las variables cualitativas o *categóricas* pueden ser *nominales* u *ordinales*, y las variables cuantitativas pueden ser *discretas* y *continuas*.

A continuación, se muestran algunos ejemplos de cada tipo de variable:

Variables cualitativas nominales:

- En un estudio sobre la biodiversidad de una reserva natural, se clasifican las especies de árboles en categorías según su tipo, como árboles de hojas caducas, árboles de hoja perenne, y árboles de coníferas. Esta clasificación es nominal, ya que no hay un orden específico entre las categorías.
- Se recopilan datos sobre los tipos de residuos encontrados en una playa durante una limpieza. Estos pueden clasificarse en categorías como “plástico”, “vidrio”, “papel” y “orgánico”, donde cada tipo de residuo es una categoría nominal sin un orden específico.

Variables cualitativas ordinales:

- Se realiza una encuesta para evaluar la satisfacción de los residentes con respecto a un proyecto de restauración de un parque. Las respuestas se clasifican en categorías como “Muy insatisfecho”, “Insatisfecho”, “Neutral”, “Satisfecho” y “Muy satisfecho”. Estas categorías tienen un orden implícito de menor a mayor satisfacción.
- Se lleva a cabo un estudio sobre la calidad del agua en diferentes ríos, y se clasifica la calidad del agua en categorías como “Muy mala”, “Mala”, “Regular”, “Buena” y “Muy buena”, según los niveles de contaminación. Estas categorías tienen un orden de menor a mayor calidad del agua.

Variables cuantitativas discretas:

- Se cuenta la cantidad de aves de diferentes especies en un área de conservación durante un día. La cantidad de aves de cada especie es una variable discreta, ya que se cuentan unidades individuales y no se pueden tener fracciones de aves.

- Se registra la cantidad de árboles de diferentes alturas en un bosque. La cantidad de árboles de cada altura es una variable discreta, ya que se cuenta el número exacto de árboles y no puede haber fracciones de árboles como un árbol y medio.



Variables cuantitativas continuas:

- Se mide la concentración de dióxido de carbono (CO_2) en diferentes áreas urbanas. La concentración de CO_2 es una variable continua, ya que puede tomar cualquier valor dentro de un rango específico, como por ejemplo entre 300 y 500 partes por millón (ppm), es decir, podría tener valores como 300,51 ppm de CO_2 .
- Se registra la temperatura del agua en un río durante un período de tiempo. La temperatura del agua es una variable continua, ya que puede variar de manera suave y continua a lo largo del tiempo, por ejemplo, de 15 °C a 20 °C, como 18,5 °C.



Continuemos con el aprendizaje mediante su participación en la actividad que se describe a continuación:



Actividad de aprendizaje recomendada

Para concluir esta semana, le invito a responder las siguientes preguntas:

- ¿Puede reconocer qué tipo de variable sería el número de árboles de Podocarpus dentro de una hectárea en el Parque Nacional Podocarpus?
- ¿Será una variable discreta o continua?

Imagine que es un guardaparque en el Parque Nacional Podocarpus y necesita realizar un estudio sobre la población de árboles de Podocarpus en una hectárea específica. Para poder analizar sus datos de manera adecuada, es fundamental que identifique correctamente si el número de árboles de Podocarpus es una variable discreta o continua. Piense

detenidamente en las características de cada tipo de variable y en cómo se presenta el conteo de árboles en el escenario planteado. Estoy seguro de que, con su conocimiento y razonamiento, podrá determinar la respuesta correcta.

Nota: por favor, complete la actividad en un cuaderno o documento Word.

Contenidos, recursos y actividades de aprendizaje recomendadas



Semana 6

Unidad 5. Introducción a la estadística descriptiva y tipos de variables

En el mundo actual, donde la información y los datos abundan, es crucial desarrollar habilidades para analizar, interpretar y comunicar los hallazgos de manera efectiva. La estadística nos brinda herramientas poderosas para organizar, resumir y visualizar datos de diversas fuentes y campos de estudio. Una de las técnicas más valiosas en la estadística es la representación gráfica. Estas nos permiten explorar patrones, tendencias y relaciones en los datos de una manera intuitiva y visualmente atractiva.

Existen diferentes tipos de gráficas diseñadas para trabajar con datos categóricos y cuantitativos.

Las gráficas de barras y de pastel son ideales para representar datos categóricos, como encuestas, preferencias o distribuciones de frecuencias.

Estas gráficas nos permiten comparar de manera sencilla las proporciones o cantidades de diferentes categorías. Por otro lado, las gráficas de líneas, de puntos y los histogramas son herramientas poderosas para visualizar datos cuantitativos, como mediciones, valores numéricos o variables continuas.

Estas gráficas nos ayudan a identificar tendencias, patrones y distribuciones en los datos.



Además de elegir el tipo de gráfica adecuada, es fundamental desarrollar un ojo crítico para interpretar correctamente la información presentada. Debemos prestar atención a los ejes, las escalas, las etiquetas y las unidades de medida, así como considerar el contexto y las posibles limitaciones o sesgos en los datos.

A lo largo de esta semana, exploraremos en profundidad cada uno de estos tipos de gráficas, sus aplicaciones y cómo interpretarlas de manera efectiva. Aprenderemos a seleccionar la representación gráfica más adecuada según el tipo de datos y el objetivo del análisis, y a comunicar nuestros hallazgos de manera clara y convincente.

Las **gráficas para datos categóricos** son representaciones visuales que muestran la distribución de datos clasificados en categorías. Permiten comparar frecuencias, proporciones o porcentajes entre diferentes grupos o categorías. En gestión ambiental, son útiles para presentar datos como tipos de ecosistemas, especies identificadas o fuentes de contaminación (Mendenhall et al., 2015).

Las gráficas para datos cuantitativos como la **gráfica de pastel** dividen un círculo en sectores proporcionales al valor de cada categoría, mientras que la **gráfica de barras** usa rectángulos de altura proporcional. Ambas son efectivas para mostrar composición y comparaciones (Douglas, 2003).

Por otro lado, la **gráfica de líneas** muestra la evolución o tendencia de una variable a lo largo del tiempo, conectando puntos de datos consecutivos con líneas. Son ideales para visualizar cambios temporales en variables ambientales (Melo et al., 2007).

La **gráfica de puntos** representa cada observación como un punto en un plano cartesiano. Son útiles para visualizar la distribución de datos y detectar patrones o valores atípicos (Gutiérrez Pulido & Vara Salazar, 2012).

Al analizar gráficas ambientales, debemos:

- Examinar escalas y unidades.

- Identificar tendencias y patrones.
- Detectar valores atípicos.
- Considerar el contexto ambiental.
- Evaluar la relevancia práctica.

Los **histogramas de frecuencia relativa** son gráficos que muestran la distribución de datos cuantitativos continuos, dividiendo el rango de datos en intervalos y mostrando la proporción de observaciones en cada intervalo. La altura de cada barra representa la frecuencia relativa de las observaciones en ese intervalo (Domínguez & Tostado, 2016).

Siga estas pautas para revisar estos temas y fortalecer sus habilidades estadísticas:

- Analice las diferencias entre gráficas de pastel y barras.
- Practique la elaboración de histogramas e interprete sus formas. Relacione la frecuencia relativa con la distribución de los datos.
- Hágase preguntas reflexivas como: ¿cómo difieren las gráficas de barras y las de pastel en su representación visual?, ¿en qué contexto utilizarías una gráfica de líneas en lugar de una de puntos?

5.5. Descripción de datos por medio de gráficas

Las gráficas desempeñan un papel fundamental en el análisis estadístico al ofrecer una representación visual de los datos, lo que facilita su interpretación y comprensión. Tanto para datos categóricos como cuantitativos, las gráficas permiten identificar patrones, tendencias y relaciones que pueden no ser evidentes al observar simplemente los números. En el caso de datos categóricos, como variables nominales u ordinales, las gráficas proporcionan una manera efectiva de visualizar la distribución de frecuencias y las proporciones de cada categoría, lo que ayuda a identificar la prevalencia de diferentes clases o grupos. Por otro lado, para datos cuantitativos, como variables continuas o discretas, las gráficas permiten visualizar la distribución de los valores, detectar la presencia de valores atípicos, comparar diferentes conjuntos de datos y evaluar la simetría, sesgo y dispersión de la distribución.

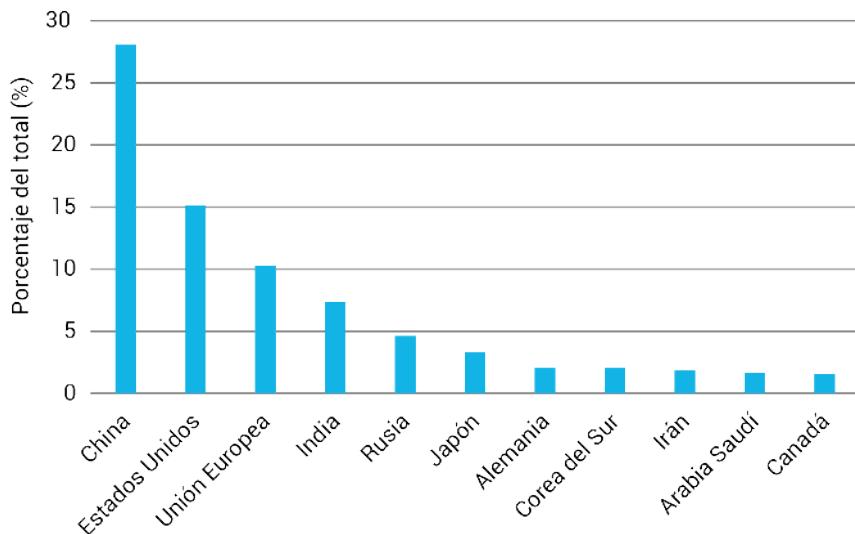
5.5.1. Gráficas para datos categóricos

Existen algunas gráficas que nos permiten representar datos categóricos de manera efectiva. La elección de la gráfica dependerá de la naturaleza de sus datos y de los patrones que deseé destacar. Para elegir la gráfica adecuada, debemos tener en cuenta el mensaje que deseamos transmitir y la audiencia o público al que vamos a dirigirnos. A continuación, se muestran en qué situaciones podríamos utilizar las gráficas para datos categóricos:

Gráfico de barras: representa las frecuencias o porcentajes de cada categoría mediante barras rectangulares. Este tipo de gráfico es adecuado para comparar la magnitud de categorías individuales y ver patrones de frecuencia. Un ejemplo sería mostrar el porcentaje de los principales países que más emisiones de CO₂ producen a nivel mundial. Para ello, le invito a revisar la siguiente figura 3.

Figura 3

Países o regiones con mayor porcentaje de emisiones de dióxido de carbono.



Nota. Adaptado de *Países más contaminantes del mundo: ranking 2022* [Ilustración], por Garrett, C., 2022, [climate.selectra](https://climate.selectra.com/), CC BY 4.0.

La figura muestra que China, Estados Unidos y la Unión Europea producen el porcentaje más alto de CO₂. Otros países o regiones como India, Rusia, Japón, Alemania, Corea del Sur, Irán, Arabia Saudita y Canadá producen porcentajes más bajos.

Gráfico de pastel (o circular): divide un círculo en secciones proporcionales a la frecuencia, proporción o porcentaje de cada categoría. Es útil para mostrar la proporción o porcentaje de cada categoría en relación con el todo. Un ejemplo sería representar el porcentaje de género en una población.

¿Cuándo utilizar un gráfico de pastel y cuándo un gráfico de barras?:

La elección entre un gráfico de pastel y un gráfico de barras depende del tipo de datos que queremos representar y de los objetivos de visualización. Utilice un gráfico de pastel cuando esté interesado en resaltar proporciones relativas y la relación entre partes y el todo. Utilice un gráfico de barras cuando necesite comparar magnitudes absolutas entre categorías.

5.5.2. Gráficas para datos cuantitativos

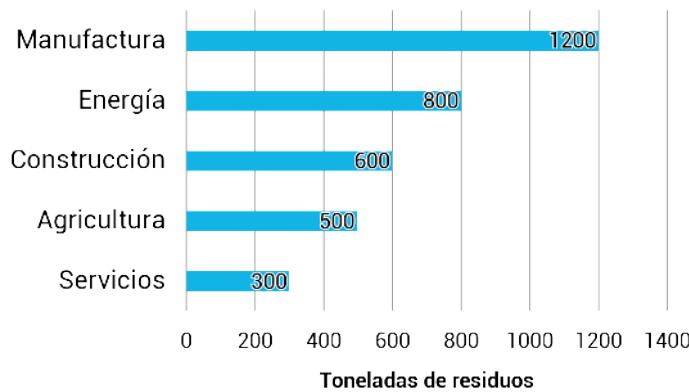
Gráficas de pastel y gráficas de barras: para representar datos cuantitativos, también podemos utilizar tanto gráficas de pastel como gráficas de barras. Imaginemos que estamos gestionando un proyecto ambiental y queremos visualizar la cantidad de residuos generados por diferentes tipos de industrias en una determinada región. Por ejemplo, tenemos datos categóricos que representan los tipos de industrias y datos continuos que indican la cantidad de residuos generados por cada industria en toneladas:

1. Agricultura: 500 toneladas de residuos.
2. Manufactura: 1200 toneladas de residuos.
3. Energía: 800 toneladas de residuos.
4. Construcción: 600 toneladas de residuos.
5. Servicios: 300 toneladas de residuos.

Antes de representar los datos cuantitativos con gráficas, es importante comprender lo siguiente. *En la gráfica de barras de la figura 4, usamos la cantidad medida (toneladas de residuos) en lugar de la frecuencia con que se presenta cada una de las categorías (tipos de industrias).* La gráfica de barras usa la altura o el largo de la barra para mostrar la cantidad de una categoría en particular. La figura 4 también podemos representarla con una gráfica de pastel.

Figura 4

Gráfica de barras horizontal para representar la cantidad de residuos generados por cada industria en toneladas.



Nota. Adaptado de *Cuáles son los sectores que más contaminan?* [Ilustración], por Marin, J., 2022, [elordenmundial](#), CC BY 4.0.

La figura muestra las toneladas de residuos generados por diferentes sectores económicos. Cada barra representa una cantidad específica de residuos en toneladas.

Gráficas de líneas: las gráficas de líneas son útiles para representar datos cuantitativos cuando queremos mostrar la relación y la tendencia entre variables continuas a lo largo de un eje x o eje horizontal, que generalmente representa el tiempo o una variable ordinal. Facilitan la identificación de patrones, tendencias y variaciones en los datos cuantitativos, haciendo que sean una herramienta valiosa en diversos contextos.

Como ejemplo, imaginemos que estamos involucrados en un proyecto de gestión ambiental que implica el monitoreo de la calidad del aire en una ciudad a lo largo de un año. Tenemos datos cuantitativos sobre la concentración de un contaminante atmosférico que produce lluvia ácida, el dióxido de azufre (SO_2):

Tabla 1

Concentración de SO_2 ($\mu\text{g}/\text{m}^3$) medidos durante un año

Mes	Concentración de SO_2 ($\mu\text{g}/\text{m}^3$)
Enero	20
Febrero	18
Marzo	15
Abril	22
Mayo	25
Junio	28
Julio	30
Agosto	27
Septiembre	24
Octubre	20
Noviembre	18
Diciembre	16

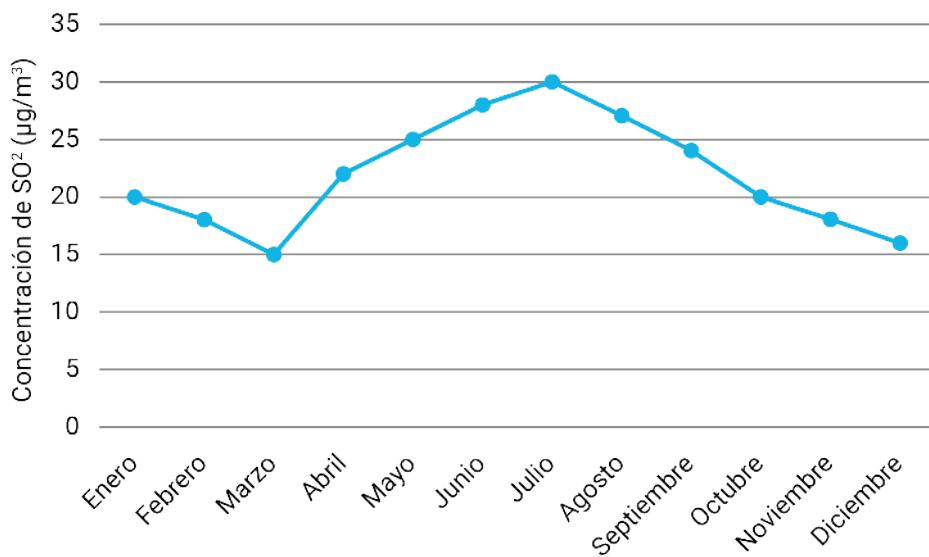
Nota. Vélez, D., 2024.

La tabla muestra la variación estacional en la concentración de SO_2 en el aire, con valores máximos en los meses de verano y mínimos hacia finales de año e inicios del año siguiente.

La gráfica sería representada de la siguiente manera:

Figura 5

Gráfica de líneas para mostrar la tendencia en la concentración de SO_2 a lo largo de un año.



Nota. Vélez, D., 2024.

Al observar la gráfica de líneas (figura 5), podemos identificar patrones y tendencias en la concentración de SO_2 a lo largo del año. Por ejemplo, ¿hay estacionalidades o picos (valores máximos) de dióxido de azufre (SO_2) en ciertos meses? La respuesta es sí, el valor máximo de dióxido de azufre ocurre en julio (ver figura 5).

5.5.3. Histogramas de frecuencia relativa

Un histograma es un tipo de gráfico utilizado para representar la distribución de frecuencias de un conjunto de datos continuos. Es de una sola variable continua. Consiste en barras rectangulares colocadas contiguamente, donde la longitud de cada barra es proporcional a la frecuencia de los datos que caen

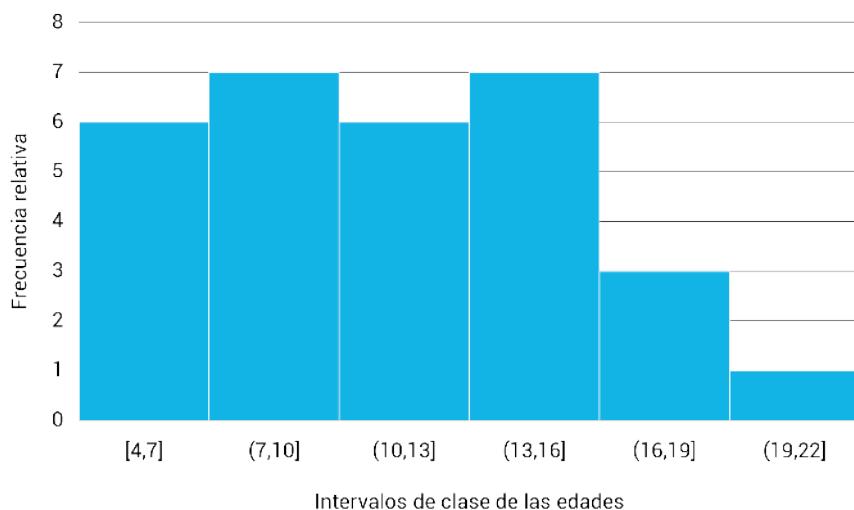
dentro de un intervalo de clase específico en el eje horizontal. En un histograma, el eje vertical representa la frecuencia o densidad de frecuencia de los datos.

La principal diferencia entre un gráfico de barras y un histograma es que el gráfico de barras se utiliza para datos categóricos (p. ej. alto, medio y bajo), mientras que el histograma se utiliza para datos continuos (p. ej. peso de una persona) o discretos (p. ej. número de árboles). El gráfico de barras representa categorías distintas, mientras que el histograma muestra la distribución de los datos en intervalos continuos. Además, en un histograma, los intervalos del eje x (o eje horizontal) son contiguos y no hay espacios entre las barras (como en la gráfica de barras), ya que representan una escala continua.

A modo de ejemplo, imaginemos que estamos trabajando en un proyecto de conservación del jaguar a nivel regional y queremos visualizar la distribución de las edades estimadas mediante cámaras trampa. Tenemos un conjunto de 30 datos que representan las edades en años de individuos dentro del grupo: 4, 8, 12, 9, 15, 7, 18, 6, 11, 14, 5, 10, 13, 16, 7, 8, 20, 9, 17, 12, 14, 6, 8, 11, 19, 10, 14, 13, 16, 15. A continuación, se muestra cómo se ve el histograma con estos datos:

Figura 6

Histograma para visualizar la distribución de las edades en un grupo de jaguares.



Nota. Vélez, D., 2024.

Este histograma proporciona una representación visual de la distribución de edades en una población determinada, con una mayor frecuencia relativa en los grupos de edad intermedios y una menor frecuencia en los grupos de mayor edad.

Al observar el histograma, podemos visualizar la distribución de las edades en el grupo de animales. Podemos identificar fácilmente que la concentración de edades está en el intervalo de 7 a 10 años (con 7 individuos), y en el intervalo de 13 a 16 años (también con 7 individuos). También podemos observar que solamente existe un individuo en el intervalo de edad de 19 a 22 años; quizás se trate de un animal ya mayor (ver figura 6).

Debemos recordar que ajustar el número de intervalos y la amplitud de estos puede influir en la apariencia del histograma, por lo que es útil experimentar con otras amplitudes de intervalos para encontrar la representación visual más informativa según los objetivos del estudio.

Después de haber estudiado esta unidad, es momento de reflexionar sobre su aprendizaje resolviendo la siguiente actividad.



Actividad de aprendizaje recomendada

Le invito a realizar la siguiente autoevaluación, la cual le permitirá identificar sus fortalezas y oportunidades para reforzar sus conocimientos.



Autoevaluación 5

1. **Ejercicio 1:** identifique cada variable como cuantitativa o cualitativa:

- Estado de conservación de una reserva natural o parque nacional.
- Concentración de dióxido de carbono (ppm) en la atmósfera.
- Nivel de conciencia ambiental para medir la percepción y el conocimiento de la población sobre temas ambientales.
- Cantidad de residuos sólidos urbanos (kg) por día en un hogar.
- Consumo de agua potable per cápita medido en litros por persona por día.

2. **Ejercicio 2:** identifique cada variable como continua o discreta:

- Temperatura del aire.
- Número de especies vegetales en un bosque.
- Nivel de ruido ambiental en una ciudad.
- Cantidad de residuos reciclados por mes medido en kg.
- Cantidad de incidentes de contaminación ambiental registrados en un año en Ecuador.

3. **Ejercicio 3:** uno de los principales contaminantes del agua de los ríos son los nitratos provenientes de los fertilizantes utilizados en la agricultura. Se hicieron mediciones de la concentración de nitratos en el agua de un río (mg/L) diariamente durante dos meses (60 datos). En la siguiente tabla se muestran los valores de los datos colectados:

15.2	13.8	14.5	14.2	13.9	15	14.4	14.7	14.9	14.1
------	------	------	------	------	----	------	------	------	------



14.6	14.3	13.7	14.8	15.1	13.5	14	13.8	14.2	14.5
14.3	14.6	13.9	15	14.4	14.7	14.9	14.1	14.6	14.3
13.7	14.8	15.1	13.5	14	13.8	14.2	14.5	14.3	14.6
13.9	15	14.4	14.7	14.9	14.1	14.6	14.3	13.7	14.8
15.1	13.5	14	13.8	14.2	14.5	17	18	17	13

- Con solo mirar la tabla, ¿qué forma piensa usted que tendrá la distribución de datos para concentración de nitratos en el agua de un río (mg/L)?
- Trace un histograma de frecuencia relativa para describir la distribución. Describa la forma de la distribución y busque resultados atípicos. ¿La gráfica confirma su respuesta al literal a?

Ir al solucionario



¡Felicitaciones por completar esta unidad! Ha adquirido habilidades valiosas en la representación gráfica de datos. ¡Ahora, prepárese para sumergirse en la próxima unidad!

Contenidos, recursos y actividades de aprendizaje recomendadas



Semana 7

La descripción de un conjunto de datos no se limita únicamente al uso de gráficas. Si bien las representaciones gráficas son herramientas poderosas para comprender patrones y tendencias en los datos, el análisis estadístico implica utilizar una combinación entre gráficas y medidas numéricas para obtener una comprensión completa de los datos.

Unidad 6. Descripción de datos con medidas numéricas

Las **medidas numéricas** complementan la visualización gráfica, proporcionando valores precisos que describen características importantes de los datos.

Los datos ambientales pueden describirse mediante:

- Medidas de tendencia central.
- Medidas de dispersión.
- Medidas de posición relativa.

Las **medidas de centro o centrales** son valores que representan el punto central o típico de un conjunto de datos como:

- a. Media aritmética: es la suma de todos los valores dividida entre el número total de observaciones. Es la medida de tendencia central más utilizada. Sin embargo, es sensible a valores extremos. Ejemplo. Si medimos el pH en 5 puntos de un río y obtenemos: 6.8, 7.2, 7.0, 6.9, 7.1. Media = $(6.8 + 7.2 + 7.0 + 6.9 + 7.1) / 5 = 7.0$. Este valor indica el nivel promedio de pH en el río.
- b. Mediana: es el valor que divide el conjunto de datos ordenados en dos partes iguales. Es especialmente útil cuando hay valores extremos o datos asimétricos. Ejemplo. Concentraciones de un contaminante (mg/L): 2.1, 2.3, 2.4, 8.9, 9.2. Mediana = 2.4. Este valor es más representativo que la media cuando hay valores atípicos como 8.9 y 9.2.
- c. Moda: es el valor que aparece con mayor frecuencia. Puede haber más de una moda o ninguna. Ejemplo. Número de especies por parcela: 5, 6, 6, 7, 6, 8, 9. Moda = 6. Indica el número de especies que se encuentra más frecuentemente.

Las **medidas de variabilidad** o dispersión indican qué tan esparcidos están los datos respecto a las medidas de centro. Son fundamentales para comprender la variación natural en datos ambientales y evaluar la confiabilidad de nuestras mediciones (Mendenhall et al., 2015).

El **rango** es una medida de variabilidad y es la diferencia entre el valor máximo y mínimo. Proporciona una idea rápida de la dispersión total de los datos. Ejemplo. Temperaturas diarias (°C): 15, 18, 20, 22, 25. Rango = 25 - 15 = 10°C. Indica la amplitud total de variación de temperatura.

La **varianza** es otra medida de variabilidad y es el promedio de los cuadrados de las desviaciones respecto a la media. Proporciona una medida de variabilidad que considera todos los datos. Ejemplo. Para concentraciones de oxígeno disuelto (mg/L): 6.2, 6.5, 6.8, 6.3, 6.6:

- Calcular la media: 6.48.
- Calcular desviaciones al cuadrado.
- Promediar las desviaciones cuadradas.
- La varianza resultante indica qué tan dispersos están los valores alrededor de la media.

La **desviación estándar** es otra medida de variabilidad y es la raíz cuadrada de la varianza. Tiene la ventaja de estar en las mismas unidades que los datos originales. Ejemplo. Siguiendo el ejemplo anterior del oxígeno disuelto, la desviación estándar sería la raíz cuadrada de la varianza, proporcionando una medida de dispersión en mg/L que es más fácil de interpretar. La elección de estas medidas dependerá de:

- Tipo de datos.
- Presencia de valores extremos.
- Objetivo del análisis.
- Audiencia del reporte.

Las **mediciones de posición relativa** nos permiten entender dónde se ubica un valor específico en relación con el conjunto completo de datos. Estas medidas son especialmente útiles para comprender la distribución de los datos y detectar valores inusuales o extremos (Mendenhall et al., 2015; Douglas, 2003).

En estudios ambientales, estas medidas nos ayudarían a:

- Establecer valores de referencia.
- Identificar niveles críticos.
- Comparar diferentes conjuntos de datos.
- Detectar valores atípicos que pueden indicar problemas ambientales.

Las principales mediciones de posición relativa son:

- Percentiles: dividen el conjunto de datos en 100 partes iguales, permitiendo determinar qué porcentaje de datos está por debajo de un valor específico. Por ejemplo, en el monitoreo de contaminantes atmosféricos, el percentil 98 se utiliza frecuentemente para evaluar el cumplimiento de estándares de calidad del aire.
- Cuartiles: son casos especiales de percentiles que dividen los datos en cuatro partes iguales, proporcionando una visión más general de la distribución. Son particularmente útiles para construir gráficas de caja y realizar comparaciones entre grupos.
- Resumen de cinco números: combina los cuartiles con los valores extremos para proporcionar una descripción completa pero concisa de la distribución de los datos. Este resumen es la base para construir las **gráficas de caja**, una herramienta visual poderosa para comparar distribuciones (Gutiérrez Pulido & Vara Salazar, 2012).

6.1. Descripción de un conjunto de datos con medidas numéricas

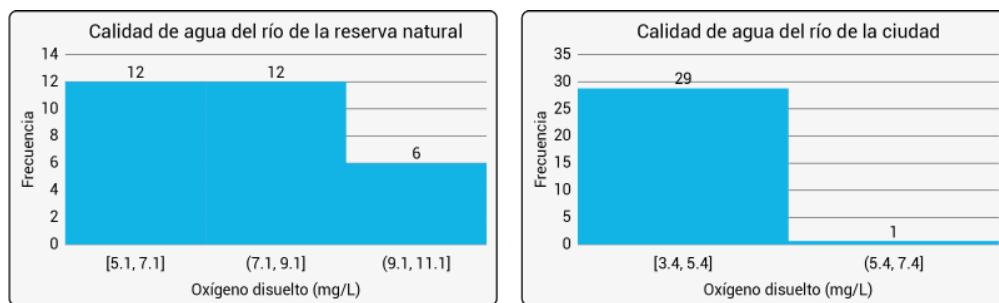
A continuación, se muestran algunos ejemplos de los conceptos revisados en esta unidad relacionados con la gestión ambiental.

Por ejemplo, supongamos que estamos trabajando en un proyecto ambiental y queremos comparar la calidad del agua de dos ríos. Un río atraviesa una ciudad y el otro atraviesa una reserva natural. Para esto, hemos medido diariamente durante un mes el oxígeno disuelto del agua de ambos ríos para

monitorear su calidad. A continuación, vamos a observar cómo las medidas numéricas y las gráficas pueden complementarse para proporcionar una mayor comprensión acerca de la calidad de agua de estos dos ríos:

Figura 7

Histograma que presenta la distribución de las mediciones sobre el oxígeno disuelto en un río que cruza una reserva natural (izquierda) y otro que atraviesa una zona urbana (derecha).



Nota. Vélez, D., 2024.

La figura de la izquierda representa un gráfico de barras que muestra la calidad del agua en términos de la concentración de oxígeno disuelto en el río de una reserva natural y la figura de la derecha muestra la calidad del agua del río en una ciudad.

Los histogramas de la figura 7 nos permiten observar la distribución de las mediciones acerca del oxígeno disuelto de ambos ríos. En el caso del agua del río que atraviesa la reserva natural, observamos seis mediciones con altos niveles de oxígeno disuelto (valores entre 9,1 y 11,1). En el caso del agua del río que atraviesa la ciudad, observamos una medición con altos niveles de oxígeno disuelto (valores entre 5,4 y 7,4). Pero, de manera general, *¿podemos decir que la calidad del agua de los dos ríos es igual o es diferente?* Una manera de responder a esta pregunta es usando *medidas numéricas* como la media (promedio), una medida de centro. El río de la reserva natural tiene una media de 7,6 mg/L de oxígeno disuelto y el río de la ciudad tiene una media de 4,4 mg/L de oxígeno disuelto. Ahora sí podemos decir que, de acuerdo con la media del oxígeno disuelto, la calidad del agua de ambos ríos es diferente.

6.2. Medidas de centro

Las medidas de centro representan la ubicación central de un conjunto de datos. Estas medidas proporcionan una indicación de dónde se concentran o agrupan los valores en la distribución de datos. Las medidas de centro son útiles porque proporcionan una forma de resumir la tendencia central de un conjunto de datos en un solo valor o un conjunto pequeño de valores. Existen varias medidas de centro, nosotros revisaremos las más utilizadas dentro de la estadística descriptiva como son la *media* o promedio, la *mediana* y la *moda*.

Cuando hablamos de medidas de centro, podemos imaginar el conjunto de datos como una línea y la medida de centro como el punto en esa línea donde la distribución parece estar más concentrada. Le propongo lo siguiente, calcule las tres medidas de centro, *media*, *mediana* y *moda*, de este conjunto de valores: 3, 1, 5, 3, 2, 6, 7, 3, 4, 8. ¿Los resultados de la media, mediana y moda están ubicados en el centro de este conjunto de datos?

6.3. Medidas de variabilidad

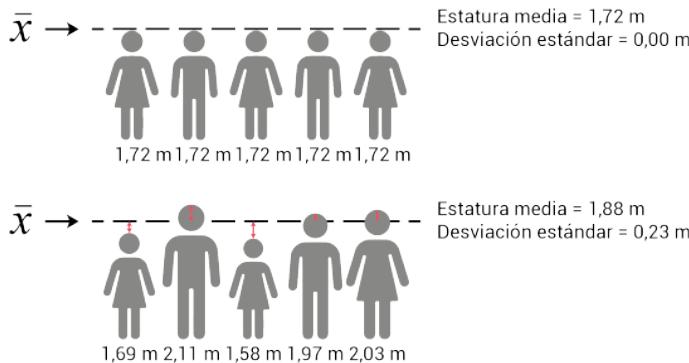
Mientras que las medidas de centro, como la media, mediana y moda, se centran en proporcionar un valor central o típico, las medidas de variabilidad se centran en indicar cuán dispersos o variables están los datos alrededor de ese valor central. Cuando pensamos en medidas de variabilidad, podemos imaginar que tan cerca o lejos están los datos alrededor de una medida central como la media. Un conjunto de datos con baja variabilidad tendrá valores cercanos a la media, mientras que un conjunto de datos con alta variabilidad mostrará valores más alejados (o dispersos) respecto a la media.

Una de las medidas de variabilidad es la *desviación estándar*. Una desviación estándar mayor indica una mayor dispersión de los datos alrededor de la media, lo que implica mayor variabilidad en el conjunto. Una desviación estándar menor sugiere que los valores están más concentrados alrededor de la media, indicando una menor variabilidad.

Por ejemplo, imaginemos un conjunto de datos de estaturas donde la media es 1,72 m y la desviación estándar es 0,00 m (ver figura 8). Significa que todas las personas tienen exactamente la misma estatura y, por tanto, no hay variabilidad en estaturas. Ahora, imaginemos un conjunto de datos de estaturas donde la media es 1,88 m y la desviación estándar es 0,23 m (figura 8). Esto significa que la mayoría de las edades están dentro de un rango de \pm 0,23 m alrededor de la media (\pm significa más/menos). Por tanto, cuanto mayor sea la desviación estándar, más dispersas o variables estarán las estaturas, y cuanto menor sea, más cercanas estarán al valor de la media.

Figura 8

Representación de la desviación estándar respecto a la media.



Nota. Vélez, D., 2024.

La figura muestra dos imágenes que representan dos distribuciones diferentes de estaturas, ofreciendo una representación visual de los conceptos de estatura promedio y desviación estándar. Estas medidas estadísticas son ampliamente empleadas para describir y comparar distribuciones de datos.

Las medidas de variabilidad son esenciales para comprender la dispersión de los datos y proporcionan información clave para la precisión de los modelos estadísticos y la toma de decisiones en campos y aplicaciones, incluido el campo de la gestión ambiental. Algunas de las medidas de variabilidad más comunes son el *rango*, la *varianza*, y la *desviación estándar*.

6.4. Mediciones de posición relativa

Las mediciones de posición relativa son herramientas valiosas en la gestión ambiental para evaluar, por ejemplo, la calidad del agua, aire o suelo, establecer límites, identificar tendencias y tomar decisiones informadas sobre la conservación del ambiente.

Por ejemplo, puede utilizar los percentiles para evaluar la calidad del agua. Como ha visto, el oxígeno disuelto es un indicador crítico de la calidad del agua, especialmente en cuerpos de agua como ríos y lagos, donde niveles inferiores a 5 mg/L pueden afectar negativamente la vida acuática (Bozorg-Haddad, et al. 2021). Va a utilizar los percentiles para evaluar si los niveles de oxígeno disuelto de un río que atraviesa la ciudad están dentro de los límites aceptables. Podría formular la pregunta, ¿qué porcentaje de los datos tiene mediciones menores a 5 mg/L de oxígeno disuelto? Esto podría responder calculando los *percentiles* de este conjunto de mediciones (tabla 2).



Tabla 2

Valores diarios de oxígeno disuelto en mg/L registrados durante un período de 30 días.

Día	Valores de oxígeno disuelto en mg/L	Día	Valores de oxígeno disuelto en mg/L	Día	Valores de oxígeno disuelto en mg/L
1	4,5	11	4	21	3,4
2	5	12	5,1	22	4,6
3	3,8	13	4,2	23	5
4	4,2	14	3,7	24	3,8
5	5,5	15	4,5	25	4,2
6	4	16	5	26	5,4
7	3,5	17	3,9	27	4
8	4,8	18	4,3	28	3,5
9	5,2	19	5,3	29	4,9
10	3,6	20	4,1	30	3,7

Nota. Vélez, D., 2024.

La tabla presenta un conjunto de mediciones diarias de oxígeno disuelto durante un mes. Los valores muestran cierta variabilidad día a día, pero en general se mantienen dentro de un rango de aproximadamente 3,4 a 5,5 mg/L.

El cálculo del percentil de un valor implica determinar en qué posición se encuentra ese valor en relación con el conjunto de datos, como las mediciones de oxígeno disuelto (Tabla 2). Antes de calcular los percentiles, debe organizar los datos en orden ascendente o descendente. Ahora, debe calcular el percentil del 75 % de acuerdo con la siguiente fórmula: $i = (P/100) * n$; donde, i representa la posición del percentil, P es el percentil que necesita calcular y n es el número de datos que tiene, en este caso son 30 mediciones de oxígeno disuelto que se tomaron diariamente por un mes (Tabla 2). Obtiene lo siguiente: $i = (75 / 100) * 30 = 22,5$. El resultado debe ser un número entero, en este caso, redondee hacia arriba y el resultado se convierte de 22,5 a **23**.

Luego, encuentre el valor en la posición calculada en el conjunto de datos ordenado de menor a mayor, como se muestra en la Tabla 3. Es importante destacar que, cuando calcule la posición de un percentil, estará determinando en qué lugar se encuentra ese percentil dentro de la secuencia de datos ordenados y, por tanto, el resultado debe ser un número entero que indique la posición exacta del percentil (ver Tabla 3).

Tabla 3

Valores de oxígeno disponible en diferentes posiciones a lo largo de un período de 30 días o mediciones.

Posición	1	2	3	4	5	6	7	8	9	10
Valores de oxígeno disponible	3,4	3,5	3,5	3,6	3,7	3,7	3,8	3,8	3,9	4,0
Posición	11	12	13	14	15	16	17	18	19	20
Valores de oxígeno disponible	4,0	4,0	4,1	4,2	4,2	4,2	4,3	4,5	4,5	4,6
Posición	21	22	23	24	25	26	27	28	29	30
Valores de oxígeno disponible	4,8	4,9	5,0	5,0	5,0	5,1	5,2	5,3	5,4	5,5

Nota. Vélez, D., 2024.

La tabla muestra un patrón claro de aumento en los valores de oxígeno disponible a medida que se avanza en las posiciones, con tres secciones distintas que presentan diferentes rangos de valores.

Esto significa que la medición de oxígeno disuelto de 5 mg/L se encuentra en el percentil 75 en relación con el conjunto de datos (Tabla 3). En otras palabras, el 75% de las mediciones de oxígeno disuelto son iguales o inferiores a 5 mg/L y el 25% de las mediciones son mayores a 5 mg/L. Un valor en el percentil 75 indica que la mayoría de las mediciones están bajo los niveles de oxígeno disuelto aceptables en comparación con las otras mediciones recopiladas. Esto podría indicar una situación preocupante, y se debería realizar la recuperación de este río para promover la oxigenación urgente de este río.

También podría utilizar otra manera de organizar un conjunto de datos para comprender cómo están agrupados o distribuidos utilizando los *cuartiles*. Los cuartiles son medidas que dividen un conjunto de datos en cuatro partes iguales, cada parte representa un 25% de los datos. Hay tres cuartiles muestrales que son comúnmente utilizados: el primer cuartil o cuartil inferior (Q_1), el segundo cuartil o mediana (Q_2) y el tercer cuartil o cuartil superior (Q_3). Es necesario recalcar que estos cuartiles son especialmente útiles cuando se trabaja con un conjunto grande de datos y pierden su utilidad cuando la cantidad de datos es pequeña.

Retomando el ejemplo del conjunto de datos de oxígeno disuelto (tabla 3), ha obtenido para este ejemplo que $Q_1 = 3,8$; $Q_2 = 4,2$ y $Q_3 = 5$. También obtuvo el *rango intercuartil* ($IQR = Q_3 - Q_1$), el cual fue de 1,2. Para estar seguro de los cálculos realizados por el programa Excel, ¿puede ayudarme a comprobar si estos cálculos son correctos? Para comprender más ampliamente el cálculo de cuartiles muestrales, le invito a revisar el video "[Medidas de tendencia central y localización](#)".

En este video podrá observar paso a paso el proceso de cálculo e interpretación de los cuartiles, lo que le ayudará a verificar los resultados obtenidos y comprender mejor su aplicación práctica.

6.5. El resumen de cinco números y la gráfica de caja

El *resumen de cinco números* es una herramienta descriptiva que proporciona información clave sobre la distribución de un conjunto de datos. Los cinco números incluyen: 1) El valor más pequeño del conjunto de datos (mínimo); 2) El valor que deja atrás al 25 % de los datos cuando están ordenados de menor a mayor (Q_1); 3) El valor que separa el conjunto de datos en dos partes iguales, con el 50 % de los datos por debajo y el 50 % por encima (la mediana o Q_2); 4) El valor que deja atrás al 75 % de los datos cuando están ordenados de menor a mayor (Q_3), y, 5) El valor más grande en el conjunto de datos (máximo). Estos cinco números proporcionan una descripción concisa de la ubicación y dispersión de los datos, y son especialmente útiles para resumir la información en conjuntos de datos grandes.

Por otro lado, la *gráfica de caja*, también conocida como *boxplot*, es la representación visual del resumen de cinco números. Sus características son:

- **Caja:** la caja representa el rango intercuartil (IQR), que es la distancia entre el primer cuartil (Q1) y el tercer cuartil (Q3). El 50 % central de los datos se encuentra dentro de la caja.
- **Línea dentro de la caja:** esta línea representa la mediana (Q2), que es el punto medio del conjunto de datos.
- **Bigotes:** los “bigotes” se extienden desde la caja hasta los valores mínimo y máximo dentro de ciertos límites.
- **Valores atípicos (*Outliers*):** son valores individuales más allá de los bigotes que se consideran posibles valores atípicos o extraños.

La gráfica de caja proporciona una visión rápida y visual de la simetría, dispersión, y la presencia de valores atípicos en un conjunto de datos. Es una herramienta efectiva para comparar distribuciones y entender la variabilidad en los datos.

Ambas herramientas, el *resumen de cinco números* y la *gráfica de caja*, son útiles para describir y visualizar la distribución de datos en estadísticas descriptivas.



A continuación, revise el siguiente video donde se explica [Cómo realizar un diagrama de cajas en Excel](#).

Es momento de aplicar sus conocimientos a través de las actividades que se han planteado a continuación:



Actividades de aprendizaje recomendadas

1. A modo de resumen, le recomiendo revisar el video “[Cuartiles introducción](#)”. En este video podrá comprender su cálculo con algunos ejemplos.



Como pudo revisar, las medidas de posición relativa, como los cuartiles, ofrecen una forma efectiva de entender la distribución de un conjunto de datos. Los cuartiles dividen un conjunto de datos en cuatro partes iguales, proporcionando información sobre la dispersión y la concentración de los valores. Estas medidas son particularmente útiles para identificar patrones y evaluar la variabilidad en una distribución.

2. Desarrolle la siguiente autoevaluación para mejorar la comprensión de esta unidad:



Autoevaluación 6

Responda verdadero o falso:

1. La descripción de un conjunto de datos se limita únicamente al uso de gráficas.
 - a. Verdadero.
 - b. Falso.

Ahora desarrolle los siguientes ejercicios:

2. **Ejercicio 1. Especies de aves** Se contó el número de especies de aves en una zona perturbada y en otra sin perturbación de un bosque tropical. En cada zona se escogieron 10 árboles, y en cada árbol se contabilizó el número de especies de aves. Aquí se presentan los dos conjuntos de datos:

Zona perturbada	Zona sin perturbación
5	4
6	7
7	8
8	9
9	10
5	6
6	8
3	6
8	7
9	9

Zona perturbada	Zona sin perturbación
4 6	10 7

- a. ¿Cuáles son la media y desviación estándar para la zona perturbada?
- b. ¿Cuáles son la media y desviación estándar para la zona sin perturbación?
- c. Compare los centros y variabilidades de las dos zonas usando los resultados de las partes a y b.
- d. Encuentre la mediana, los cuartiles superior e inferior y el IQR para cada uno de los dos conjuntos de datos.
- e. Construya dos gráficas de caja para comparar los dos conjuntos de datos.
- f. ¿Qué conclusiones puede extraer sobre la variación del número de especies de aves en las dos zonas, considerando que cada árbol alberga un número diferente de especies de aves?

3. **Ejercicio 2. Agua de mar contaminada** La contaminación causada por el petróleo en mares y océanos estimula el crecimiento de algunos tipos de bacterias. Un conteo de microorganismos que se originan en el petróleo (bacterias por 100 mililitros) en 10 partes de agua de mar dieron estas lecturas: 49, 70, 54, 67, 59, 40, 61, 69, 71, 52

- a. Calcule el valor de s usando la aproximación de rango.
- b. Calcule \bar{x} y s y compare con la aproximación de rango de la parte a.
- c. Construya una gráfica de caja para los datos y úsela para describir la distribución de datos.

[Ir al solucionario](#)



¡Felicitaciones por dominar la descripción de datos con medidas numéricas! Está listo para sumergirse en la próxima unidad.



Resultado de aprendizaje 1 a 4:

- Conoce y aplica los pasos para el diseño de investigación.
- Realiza revisiones bibliográficas efectivas.
- Plantea hipótesis en torno al tema de investigación.
- Levanta, analiza e interpreta información de campo.

Contenidos, recursos y actividades de aprendizaje recomendadas



Semana 8



Actividades de aprendizaje recomendadas

Estrategia de estudio para el primer bimestre:

Estimado estudiante, revise a continuación una estrategia de estudio para la evaluación del primer bimestre.

1. Revisión de apuntes y materiales:

- Dedíquese a revisar todos sus apuntes, resúmenes y materiales de estudio de las semanas 1 a 7.
- Identifique los conceptos clave, definiciones importantes y fórmulas destacadas.
- Cree una lista de los temas principales abordados en cada unidad.

2. Creación de un esquema integrador:

- Cree un esquema o mapa conceptual que integre los contenidos de las diferentes unidades.
- Establece conexiones entre los conceptos, mostrando cómo se relacionan y complementan entre sí.
- Este esquema le servirá como una guía visual para repasar los temas de manera estructurada.



3. Resolución de ejercicios y problemas clave:

- Resuelva ejercicios y problemas representativos de cada unidad.
- Concéntrese en aquellos que abarquen los conceptos fundamentales y las habilidades necesarias para la evaluación.
- Revise las soluciones y asegúrese de comprender cada paso del proceso.



4. Enseñanza a un compañero o autoevaluación:

- Reúnase con un compañero de clase y turnense para explicar los conceptos y resolver problemas.
- Si estudia solo, explique los temas en voz alta como si estuviera enseñándole a alguien más.
- Esta práctica le ayudará a consolidar su comprensión e identificar áreas que necesitan más atención.



5. Simulacro de evaluación:

- Realice un simulacro de evaluación con preguntas y problemas similares a los que espera encontrar en la evaluación real.
- Establezca un límite de tiempo y trabaje en condiciones lo más parecidas posible a las del examen.
- Revise sus respuestas y analice su desempeño, identificando áreas de mejora.

6. Repaso final y aclaración de dudas:

- Finalmente, realice un repaso final de los conceptos y temas clave.
- Revise nuevamente su esquema integrador y los ejercicios resueltos.
- Aclare cualquier duda restante con la ayuda de sus apuntes, libros o el profesor.

Durante toda la octava semana, asegúrese de mantener sesiones de estudio focalizadas y, tome descansos regulares para mantener su mente fresca y receptiva. Duerma lo suficiente, aliméntese bien y mantenga una actitud positiva. Recuerde que el objetivo es consolidar su comprensión de los temas y llegar a la evaluación con confianza en sus habilidades.



Si sigue esta estrategia de repaso intensivo durante la semana previa a la evaluación, estará bien preparado para demostrar su dominio de los contenidos estudiados en las semanas 1 a 7.





Segundo bimestre

Resultado de aprendizaje 4:

Levanta, analiza e interpreta información de campo.

La finalidad del resultado de aprendizaje es aprender cómo clasificar variables, cómo usar gráficas y medidas numéricas para describir y construir una imagen mental de los datos. Además, el estudiante aprenderá las técnicas que se pueden emplear al describir estas variables para identificar posibles relaciones entre variables.

Contenidos, recursos y actividades de aprendizaje recomendadas

Recuerde revisar de manera paralela los contenidos con las actividades de aprendizaje recomendadas y actividades de aprendizaje evaluadas.



Semana 9

Unidad 7. Descripción de datos bivariados

En gestión ambiental, frecuentemente necesitamos estudiar cómo se relacionan dos variables entre sí. Los datos bivariados nos permiten examinar estas relaciones y comprender mejor los fenómenos ambientales complejos. Por ejemplo, podemos estudiar cómo la temperatura del agua afecta la concentración de oxígeno disuelto, o cómo los niveles de contaminación se relacionan con la biodiversidad de un ecosistema.

El análisis de datos bivariados incluye tanto métodos gráficos como numéricos. Las representaciones gráficas nos permiten visualizar patrones y relaciones, mientras que las medidas numéricas nos ayudan a cuantificar la fuerza y dirección de estas relaciones. Esta combinación de herramientas es esencial para tomar decisiones informadas en la gestión y conservación ambiental.

A continuación, exploraremos diferentes métodos para analizar y representar datos bivariados, enfocándonos en su aplicación práctica en estudios ambientales.

Los **datos bivariados** son observaciones donde se miden dos variables diferentes en cada unidad de estudio. Estas variables pueden ser cualitativas, cuantitativas o una combinación de ambas. En estudios ambientales, los datos bivariados son fundamentales para entender las relaciones entre diferentes parámetros ambientales (Mendenhall et al., 2015).

Cuando trabajamos con variables cualitativas, podemos utilizar:

- a. Gráficas de barras agrupadas: permiten comparar frecuencias entre diferentes categorías. Estas muestran relaciones entre variables categóricas. Ejemplo. Comparar la abundancia de diferentes especies en distintos tipos de hábitat.
- b. Gráficas de barras apiladas: estas muestran la composición total y proporcional. Son útiles para visualizar distribuciones conjuntas. Ejemplo. Distribución de tipos de vegetación en diferentes zonas altitudinales.
- c. Gráficas de dispersión: son fundamentales para visualizar relaciones entre dos variables cuantitativas. En estas gráficas cada punto representa un par de valores. El patrón de puntos sugiere el tipo de relación. Las gráficas de dispersión permiten identificar tendencias y valores atípicos. Ejemplo. Un gráfico que muestre la relación entre:
 - Temperatura del agua (eje X) y oxígeno disuelto (eje Y).
 - Precipitación (eje X) y caudal del río (eje Y).

Las **medidas numéricas para datos cuantitativos bivariados** son herramientas estadísticas que nos permiten cuantificar y describir la relación entre dos variables numéricas. Estas medidas complementan el análisis gráfico y proporcionan valores específicos que indican tanto la fuerza como la dirección de la relación entre variables (Mendenhall et al., 2015).

En estudios ambientales, estas medidas son particularmente útiles para:

- Cuantificar relaciones entre variables ambientales.
- Evaluar la fuerza de asociaciones.
- Predecir comportamientos.
- Tomar decisiones basadas en la evidencia.

Las dos medidas principales son:

- Covarianza: la covarianza es una medida que indica cómo dos variables cambian juntas. Un valor positivo indica que las variables tienden a aumentar o disminuir juntas, mientras que un valor negativo indica que cuando una variable aumenta, la otra tiende a disminuir. Sin embargo, su interpretación puede ser difícil porque depende de las unidades de medida de las variables (Melo et al., 2007).
- Coeficiente de correlación: el coeficiente de correlación es una versión estandarizada de la covarianza que varía entre -1 y +1. Esta escala estandarizada facilita la interpretación y permite comparar relaciones entre diferentes pares de variables, independientemente de sus unidades de medida. Un valor cercano a +1 indica una fuerte relación positiva, mientras que un valor cercano a -1 indica una fuerte relación negativa (Gutiérrez Pulido & Vara Salazar, 2012).

Para comprender los contenidos indicados, hágase preguntas reflexivas como:

- ¿Cuáles son las ventajas de utilizar gráficas de barras en comparación con gráficas circulares?
- ¿Puede pensar en situaciones en las que un gráfico de barras apiladas sea más efectivo que uno agrupado?
- ¿Cuál es el propósito principal de una gráfica de dispersión?
- ¿Cómo interpretarías visualmente la relación entre dos variables cuantitativas en una gráfica de dispersión?
- ¿Cómo interpretarías visualmente la relación entre dos o más categorías de una variable cualitativa en una gráfica de cajas?

A continuación, se muestran algunos ejemplos de los conceptos revisados en esta unidad relacionados con la gestión ambiental.

7.1. Datos bivariados

Los datos bivariados involucran dos variables, permitiéndole explorar la relación entre ellas. La relación entre una variable independiente y una variable dependiente es uno de los enfoques más importantes de la investigación científica. Cada observación tiene valores para ambas variables.

Por ejemplo, considere que ha recopilado datos de 25 áreas urbanas diferentes, donde ha medido tanto la densidad del tráfico (vehículos por hora, variable 1) como la concentración de dióxido de nitrógeno (NO_2 , variable 2). Cada observación representa un par de valores únicos para una ubicación específica. El análisis de datos bivariados es crucial en diversas disciplinas, como la economía, la psicología, la gestión ambiental y muchas otras, donde se busca entender la relación entre diferentes variables.

7.2. Gráficas para variables cualitativas

Las gráficas para variables cualitativas se utilizan para representar la distribución de datos categorizados o cualitativos. Estas gráficas son especialmente útiles para comprender la frecuencia, proporción o relación entre diferentes categorías.

Por ejemplo, su objetivo es evaluar la calidad del agua de ríos urbanos y no urbanos en las tres regiones naturales de Ecuador mediante la medición de la riqueza de especies de macroinvertebrados bentónicos (riqueza es el número de especies). La variable cualitativa en este caso son las regiones naturales de Ecuador: costa, sierra y oriente. La riqueza de especies de macroinvertebrados es la variable dependiente. Los datos de la tabla 4 son datos ficticios con fines educativos y representan una muestra de cinco ríos urbanos y cinco ríos no urbanos en cada región, en los cuales constan el número de especies de macroinvertebrados.



Tabla 4

El valor en cada celda es el número promedio de especies de macroinvertebrados para los ríos que cayeron en esa categoría.

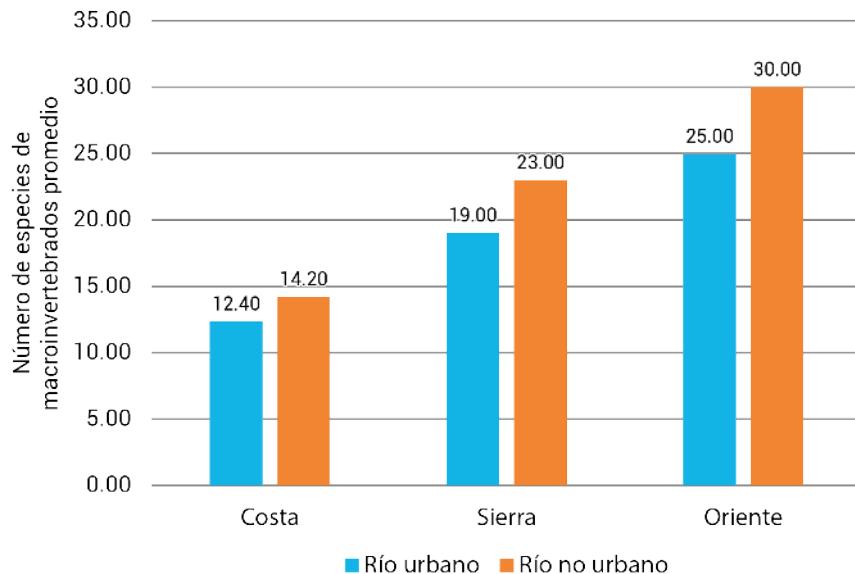
	Costa	Sierra	Oriente
Río urbano	12.40	19.00	25.00
Río no urbano	14.20	23.00	30.00

Nota. Vélez, D., 2024.

Para presentar el número de especies de macroinvertebrados de cada río, se puede usar una gráfica de barras lado a lado, como se muestra en la figura 9. Puede observar que los ríos no urbanos del Oriente tienen el número promedio más alto de especies de macroinvertebrados en comparación al resto de regiones.

Figura 9

Gráficas de barras para visualizar el número de especies de invertebrados en ríos urbanos y no urbanos de las tres regiones naturales de Ecuador.



Nota. Vélez, D., 2024.

La figura es un gráfico de barras agrupadas que compara el número de especies de macroinvertebrados promedio en ríos urbanos y no urbanos (o rurales) en tres regiones diferentes: Costa, Sierra y Oriente.

7.3. Gráficas de dispersión para dos variables cuantitativas

Las gráficas de dispersión sirven para representar la relación entre dos variables cuantitativas. En estas gráficas, cada punto en el plano cartesiano representa una observación que tiene valores específicos para ambas variables. Son útiles para visualizar patrones, tendencias y la fuerza de la relación entre las dos variables.

Por ejemplo, en la tabla 5 se muestran dos conjuntos de datos ficticios con fines educativos, uno para oxígeno disuelto (medido en mg/L) y otro para la abundancia de macroinvertebrados bentónicos. Es importante entender cuál es la variable independiente y cuál es la variable dependiente para poder interpretar la relación entre estas dos variables correctamente. En este caso, el oxígeno disuelto es la variable que se presume causa un efecto sobre la abundancia de macroinvertebrados (no al revés).

Tabla 5

Relación entre la concentración de oxígeno disuelto (en mg/L) y la abundancia de macroinvertebrados en un ecosistema acuático.

Oxígeno disuelto (x) (mg/L)	Abundancia de macroinvertebrados (y)
7.2	45
6.8	38
7.0	50
6.5	32
6.9	42
7.1	47
6.6	35
6.3	28
7.3	53
6.7	40

Nota. Vélez, D., 2024.

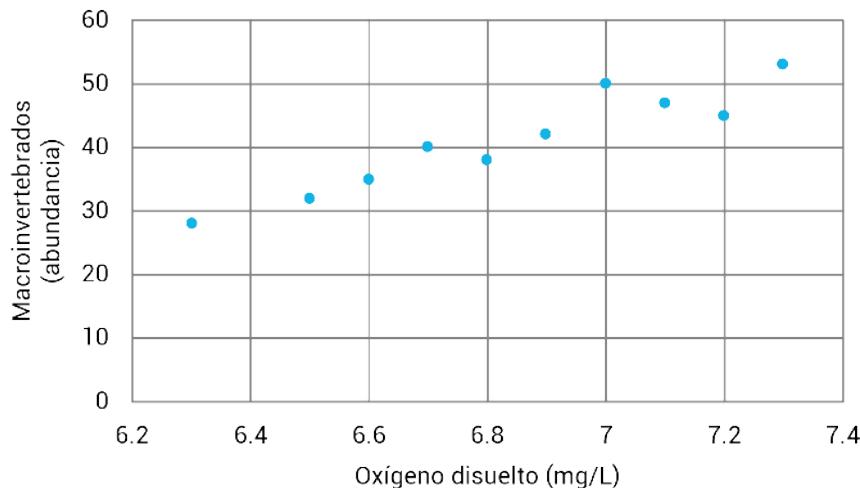
La tabla muestra un conjunto de datos que relacionan el oxígeno disuelto con la abundancia de macroinvertebrados en un sistema acuático. Aunque se observan algunas variaciones en la abundancia a diferentes niveles de oxígeno, no hay un patrón consistente.

Con la ayuda de una gráfica de dispersión, puede observar una relación y tendencia entre las dos variables. La abundancia de macroinvertebrados (y) aumenta con el oxígeno disuelto en una tendencia constante hacia arriba (figura 10).



Figura 10

Diagrama de dispersión para mostrar la relación y tendencia entre dos variables. La abundancia de macroinvertebrados (*y*) aumenta con el oxígeno disuelto en una tendencia constante hacia arriba.



Nota. Vélez, D., 2024.

La figura es un gráfico de dispersión que muestra la relación entre el oxígeno disuelto (en mg/L) en el eje horizontal y la variable "macroinvertebrados (abundancia)" en el eje vertical.

7.4. Medidas numéricas para datos cuantitativos bivariados

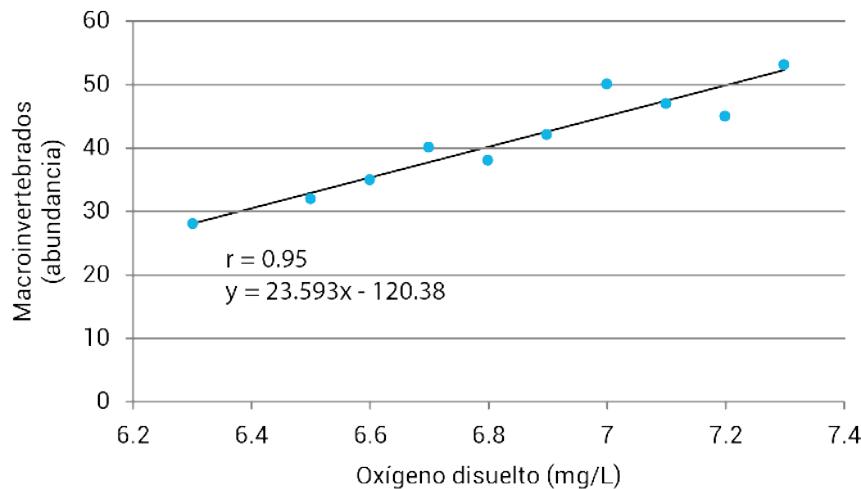
Las medidas numéricas para datos cuantitativos bivariados se utilizan para resumir y describir la relación entre dos variables cuantitativas. Estas medidas proporcionan información cuantitativa sobre cómo las variables están relacionadas entre sí.

Por ejemplo, continuando con el ejemplo anterior, la gráfica de dispersión de la figura 11 muestra un modelo lineal, es decir, una recta con los puntos de datos arriba y debajo de la recta (línea negra). En este caso, decimos que la abundancia de macroinvertebrados y el oxígeno disuelto exhiben una *relación lineal positiva*. En la figura también se indica el valor del coeficiente de correlación (*r*), el cual es bastante cercano a 1, lo cual indica que la relación

lineal entre las dos variables es muy fuerte. Los cálculos matemáticos e interpretación de la relación lineal se tratan con mayor profundidad en la asignatura de *Estadística Inferencial*.

Figura 11

Diagrama de dispersión con una línea de tendencia lineal. También se muestra el valor del coeficiente de correlación ($r = 0,95$).



Nota. Vélez, D., 2024.

La figura muestra un gráfico de dispersión con una línea de regresión lineal ajustada a los datos. La variable en el eje horizontal es el nivel de oxígeno disuelto (mg/L), mientras que la variable en el eje vertical es la abundancia de macroinvertebrados. El coeficiente de correlación $r = 0.95$ indica una correlación positiva muy fuerte entre las dos variables.

Es momento de aplicar sus conocimientos a través de las actividades que se han planteado a continuación:



Actividades de aprendizaje recomendadas

1. Para comprender cómo realizar el diagrama de dispersión y cómo calcular el coeficiente de correlación en Excel, revise el video: "[Tutorial gráfico de dispersión y correlación en Excel fácil!](#)".

El diagrama de dispersión es una representación gráfica que ilustra la relación entre dos variables en un conjunto de datos. Cada punto en el gráfico representa una observación, permitiendo visualizar patrones y tendencias. La correlación, medida mediante coeficientes como Pearson, evalúa la fuerza y dirección de la relación entre las variables.

Un coeficiente cercano a 1 indica una correlación positiva, mientras que cerca de -1 indica una correlación negativa. Un coeficiente cercano a 0 sugiere una relación débil. El diagrama de dispersión y la correlación son herramientas esenciales para comprender la asociación entre variables cuantitativas y facilitar la identificación de patrones en los datos.

2. Para comprobar la comprensión de los conceptos de esta unidad, analice y resuelva la siguiente autoevaluación.



Autoevaluación 7

Responda verdadero o falso:

1. Los datos bivariados involucran dos variables, permitiendo explorar la relación entre ellas.
 - a. Verdadero.
 - b. Falso.
2. Las gráficas para variables cualitativas se utilizan para representar la distribución de datos categorizados o cualitativos.
 - a. Verdadero.
 - b. Falso.
3. En las gráficas de dispersión, cada punto en el plano cartesiano representa una observación con valores específicos para solo una variable.
 - a. Verdadero.

- b. Falso.
4. Las medidas numéricas para datos cuantitativos bivariados solo proporcionan información cualitativa sobre cómo las variables están relacionadas entre sí.

- a. Verdadero.
b. Falso.

Lo invito a desarrollar el siguiente ejercicio:

5. Ejercicio 1: la generación centennials se caracteriza por sus hábitos de consumo, comportamiento social y digital, así como por sus filosofías y tendencias de comportamiento. Esta generación aparentemente también se preocupa por su aspecto físico. A continuación, se presenta información de 10 jóvenes centennials con algunos datos.

Uso de tecnología (horas al día)	6	7	5	8	4	6	7	5	6	8
Dieta (1=Omnívoro, 2=Vegetariano, 3=Vegano)	1	2	1	1	3	2	1	2	1	1
Horas de ejercicio por semana	3	4	2	5	3	2	4	3	5	4
Edad (años)	21	23	22	24	20	23	22	25	20	24
Peso (kg)	65	70	60	75	55	68	72	67	58	71

- a. ¿Qué pares de variables espera usted que estén fuertemente relacionados?
b. Trace una gráfica de dispersión para uso de tecnología y peso. Describa la relación.

- c. Trace una gráfica de dispersión para uso de tecnología y horas de ejercicio a la semana. Compare el modelo con el que se encontró en la parte b.
- d. Trace una gráfica de dispersión para horas de ejercicio contra peso y otra para edad contra peso. Compare los modelos. ¿Hay conglomerados o resultados atípicos?
- e. Para los pares de variables que parecen estar linealmente relacionados, calcule en Excel utilizando la función "COEF. DE.CORREL" los coeficientes de correlación.
- f. Escriba un párrafo para resumir las relaciones que usted pueda ver en estos datos. Use las correlaciones y los modelos de las cuatro gráficas de dispersión para verificar sus conclusiones.

[Ir al solucionario](#)



¡Felicitaciones por terminar el estudio de esta unidad! Ahora, prepárese para sumergirse en el mundo de las probabilidades.

Resultado de aprendizaje 5:

Genera investigaciones científicas y aplica sus resultados como aporte a la solución de problemas ambientales.

La finalidad del resultado de aprendizaje es comprender los principios de la probabilidad y cómo aplicar distribuciones de probabilidad. Estos conocimientos le permitirán evaluar y prever eventos ambientales, fundamentando sus decisiones en datos y contribuyendo así a la resolución de problemas ambientales.

Contenidos, recursos y actividades de aprendizaje recomendadas

Recuerde revisar de manera paralela los contenidos con las actividades de aprendizaje recomendadas y actividades de aprendizaje evaluadas.



Semana 10

Unidad 8. Probabilidad y distribuciones de probabilidad

La probabilidad desempeña un papel fundamental en estadística, permitiéndole cuantificar la incertidumbre y tomar decisiones en situaciones aleatorias. En esta unidad explorará el papel que desempeña la probabilidad en el campo de la estadística y cómo se convierte en una herramienta esencial para entender y analizar datos. Prepárese para este viaje, liberando el poder de anticipar y comprender el azar en la toma de decisiones. ¡Vamos a empezar!

8.1. El papel de la probabilidad en estadística

Imagine un escenario en el que usted es el encargado de la Unidad de Gestión Ambiental de un municipio y tiene como objetivo, diseñar un plan eficiente para la recolección de residuos, considerando factores como la cantidad diaria de desechos generados en diferentes áreas de la ciudad. Las probabilidades en estadística pueden desempeñar un papel crucial en este proceso. Por



ejemplo, utilizando datos históricos de generación de residuos, se pueden aplicar técnicas de probabilidad para prever períodos de alta generación de residuos, como festividades o eventos especiales.

Esto permite una planificación proactiva para hacer frente a aumentos temporales en la producción de desechos.

La **probabilidad** es una herramienta matemática fundamental que permite cuantificar la incertidumbre y evaluar la posibilidad de ocurrencia de eventos en situaciones donde el resultado no puede predecirse con certeza absoluta, es decir, no es determinístico (un resultado es determinístico cuando, bajo las mismas condiciones iniciales, siempre se obtiene el mismo resultado) (Gavilánez Luna, 2021). En estadística, la probabilidad cumple dos funciones esenciales:

1. Proporciona modelos teóricos para describir la variabilidad inherente en procesos naturales y antropogénicos.
2. Permite evaluar la confiabilidad de las conclusiones extraídas de datos muestrales para hacer inferencias sobre poblaciones.

La teoría de la probabilidad establece el marco matemático necesario para (Gutiérrez Pulido & Vara Salazar, 2012)

- Modelar fenómenos aleatorios: como la dispersión de contaminantes atmosféricos en función de variables meteorológicas (dirección del viento, temperatura, presión atmosférica).
- Cuantificar la incertidumbre en mediciones y observaciones: como el margen de error en la medición de la concentración de metales pesados en muestras de agua de un río.
- Evaluar riesgos y tomar decisiones informadas: como determinar la probabilidad de que una especie se extinga en los próximos 50 años bajo diferentes escenarios de cambio climático.
- Realizar inferencias estadísticas válidas: como estimar la población total de una especie en peligro de extinción a partir de conteos realizados en áreas de muestreo.

Los **eventos** y el **espacio muestral** son conceptos fundamentales en teoría de probabilidad. El espacio muestral (S) es el conjunto de todos los posibles resultados de un experimento o proceso de observación (Melo et al., 2007). Este puede ser:

- Finito: contiene un número limitado de resultados posibles. Por ejemplo, al clasificar el estado de conservación de una especie según la UICN, los posibles resultados son: Extinto (EX), Extinto en Estado Silvestre (EW), En Peligro Crítico (CR), En Peligro (EN), Vulnerable (VU), Casi Amenazado (NT), Preocupación Menor (LC).
- Infinito numerable: los resultados pueden contarse, pero son infinitos. Por ejemplo, el número de individuos de una especie que pueden observarse en un área protegida: 0, 1, 2, 3, ... Hasta el infinito. Aunque teóricamente la secuencia continúa infinitamente, cada número puede ser identificado y contado.
- Infinito no numerable: contiene un continuo de resultados posibles. Por ejemplo, la medición de la temperatura exacta de un río en grados Celsius podría ser cualquier valor dentro de un rango: 20.1 °C, 20.11 °C, 20.111 °C, ... Entre dos valores cualesquiera siempre existe otro valor posible, formando un continuo de posibilidades.

Un **evento** es un subconjunto del espacio muestral y representa un resultado o conjunto de resultados de interés. Los eventos pueden clasificarse en:

1. Eventos simples: contienen un único resultado elemental. Por ejemplo, que una especie sea clasificada específicamente como "En Peligro (EN)" según la UICN.
2. Eventos compuestos: Formados por múltiples eventos simples. Por ejemplo, que una especie se encuentre en alguna categoría de amenaza, lo cual incluye: "En Peligro Crítico (CR)", "En Peligro (EN)" o "Vulnerable (VU)".
3. Eventos mutuamente excluyentes: No pueden ocurrir simultáneamente. Por ejemplo, una especie no puede estar clasificada como "En Peligro (EN)" y "Preocupación Menor (LC)" al mismo tiempo.
4. Eventos complementarios: su unión constituye el espacio muestral completo. Por ejemplo, una especie está "En Peligro (EN)" o "No está En

Peligro" (todas las demás categorías). Estos dos eventos juntos abarcan todas las posibilidades del espacio muestral.

La **probabilidad de un evento** se define como una medida numérica de la posibilidad de su ocurrencia, que debe satisfacer tres axiomas fundamentales (Melo et al., 2007):

1. No negatividad: la probabilidad de cualquier evento debe ser mayor o igual a cero. Por ejemplo, la probabilidad de que una especie esté "En Peligro (EN)" no puede ser -0.2; debe ser un valor entre 0 y 1.
2. Normalización: la probabilidad del espacio muestral completo es igual a uno. Por ejemplo, la suma de las probabilidades de todas las categorías posibles de estado de conservación (EX, EW, CR, EN, VU, NT, LC) debe ser igual a 1 o 100 %.
3. Aditividad: para eventos mutuamente excluyentes, la probabilidad de su unión es la suma de sus probabilidades individuales. Por ejemplo, si:

- $P(CR) = 0.2$ (probabilidad de que una especie esté "En Peligro Crítico").
- $P(EN) = 0.3$ (probabilidad de que una especie esté "En Peligro").
- $P(VU) = 0.1$ (probabilidad de que una especie esté "Vulnerable").
- Entonces, la probabilidad de que una especie esté en alguna categoría de amenaza es: $P(CR \cup EN \cup VU) = P(CR) + P(EN) + P(VU) = 0.2 + 0.3 + 0.1 = 0.6$ o 60%.

Para el cálculo de probabilidades de eventos sencillos, se pueden utilizar diferentes enfoques:

- a. Enfoque clásico (a priori): basado en la equiprobabilidad de resultados elementales. Por ejemplo, si en una reserva natural hay 100 especies de aves registradas, de las cuales 20 son rapaces, la probabilidad de seleccionar al azar una especie rapaz sería: $P(\text{rapaz}) = 20/100 = 0.2$ o 20 %.
- b. Enfoque frecuentista (a posteriori): Basado en la frecuencia relativa de ocurrencia en múltiples repeticiones [$P(A) = \lim(n \rightarrow \infty)$ (frecuencia de A)/(número total de repeticiones)]. Por ejemplo, si durante 1000 muestreos de biodiversidad en un bosque, se ha detectado la presencia de una especie



particular en 300 ocasiones, la probabilidad estimada de encontrar esta especie en un muestreo sería: $P(\text{especie presente}) = 300/1000 = 0,3$ o 30%.

c. Enfoque subjetivo: basado en el grado de creencia personal sobre la ocurrencia del evento, en el cual se debe actualizar con nueva información mediante el teorema de Bayes. Por ejemplo, un experto podría asignar inicialmente una probabilidad de 0,7 a que una especie sobreviva los próximos 10 años basándose en su experiencia. Esta probabilidad se actualizaría al obtener nuevos datos sobre la población, amenazas o éxito reproductivo.

Las propiedades fundamentales se pueden ejemplificar así:

1. $0 \leq P(A) \leq 1$: la probabilidad de encontrar una especie en peligro de extinción debe estar entre 0 y 1. No puede ser -0,3 ni 1,5.
2. $P(S) = 1$: la suma de las probabilidades de todos los posibles estados de conservación de una especie debe ser 1 (100 %).
3. $P(A \cup B) = P(A) + P(B)$: Si $P(\text{CR}) = 0,2$ y $P(\text{EN}) = 0,3$, entonces la probabilidad de que una especie esté en alguna de estas dos categorías es $0,2 + 0,3 = 0,5$.
4. $P(A') = 1 - P(A)$: Si la probabilidad de que una especie esté en peligro es 0,3, entonces la probabilidad de que no esté en peligro es $1 - 0,3 = 0,7$.

Para la comprensión de estos contenidos le recomiendo el siguiente método de estudio:

- Identifique los conceptos clave y los objetivos del estudio de la probabilidad en este contexto.
- Enfóquese en las definiciones de eventos y espacio muestral. Asegúrese de comprender claramente qué constituye un evento y cómo se define el espacio muestral en el contexto de la probabilidad.
- Intente relacionar el concepto de probabilidad con situaciones del mundo real.
- Utilice diagramas y representaciones visuales para entender mejor la relación entre eventos y espacio muestral.

- Practique con ejercicios que involucren la identificación de eventos y espacio muestral.
- Revise los ejemplos que involucren el cálculo de probabilidades con eventos sencillos.
- También puede revisar las actividades de aprendizaje recomendadas de esta semana y resolver la autoevaluación al finalizar esta unidad (semana 12).

8.2. Eventos y el espacio muestral



Imagine que está estudiando el crecimiento poblacional y la generación de residuos en una ciudad durante un año. Este proceso se puede ver como un experimento aleatorio, ya que factores como tasas de natalidad, migración y políticas de gestión de residuos hacen que los resultados no sean predecibles con certeza.

El *espacio muestral*, en este caso, sería el conjunto de todos los posibles resultados al final del año. Incluiría diferentes escenarios de crecimiento poblacional y niveles de generación de residuos, teniendo en cuenta la variabilidad en las tasas de natalidad, migración, políticas urbanas y otras variables.

En el caso de los eventos, por ejemplo, podría tener un evento que represente un crecimiento poblacional superior al promedio y una generación de residuos que excede cierto límite establecido por las normativas ambientales. Otro evento podría ser un escenario en el que la población disminuye, pero la generación de residuos per cápita aumenta.



8.3. Cálculo de probabilidades con el uso de eventos sencillos

Continuando con el ejemplo de la sección anterior (8.2), la probabilidad en este contexto se calcularía evaluando la frecuencia relativa de los eventos de interés en relación con el espacio muestral total. Por ejemplo, podría calcular la probabilidad de que la ciudad experimente un crecimiento poblacional moderado y una generación de residuos sostenibles.

Este enfoque podría proporcionar una manera de abordar la complejidad del crecimiento poblacional y la generación de residuos en una ciudad, permitiendo la aplicación de conceptos estadísticos para comprender y prever posibles escenarios futuros.

Para profundizar en el tema, le animo a revisar la siguiente presentación de diapositivas (PPT) titulada [Cálculo de probabilidades](#).

Luego de revisar la presentación, lo invito a revisar el siguiente *quiz*: “Cálculo de probabilidades”, el cual consta de dos ejercicios que contribuirán a consolidar lo aprendido.

[Cálculo de probabilidades](#)

Es momento de aplicar sus conocimientos a través de la actividad que se ha planteado a continuación.



Actividad de aprendizaje recomendada

Para comprender qué es un experimento aleatorio, espacio muestral y evento, revise el video: “[Experimento aleatorio, espacio muestral, evento o suceso y probabilidades](#)”.

Como pudo ver en el video, un experimento aleatorio es un proceso en el cual se obtiene una observación o medición, mientras que el espacio muestral representa el conjunto de todos los posibles resultados de ese experimento. Además, un evento es un conjunto específico de resultados dentro del espacio muestral. Puede consistir en un solo resultado o en

varios resultados posibles. Los eventos son las situaciones o condiciones particulares en las que estamos interesados en analizar o medir la probabilidad.



Siga adelante con esa misma dedicación y atención al detalle, ya que estos conceptos son fundamentales para construir una base sólida en estadística. ¡Bien hecho!



Contenidos, recursos y actividades de aprendizaje recomendadas



Semana 11

Unidad 8. Probabilidad y distribuciones de probabilidad

8.4. Relaciones de evento y reglas de probabilidad

Puede explorar más a fondo el ejemplo de la semana anterior relacionado con el crecimiento poblacional y la generación de residuos en una ciudad para comprender los conceptos de relaciones de evento y reglas de probabilidad.

Imagine una ciudad en la que quiere estudiar dos eventos relacionados: Evento A, que representa un aumento en la población, y Evento B, que representa un aumento en la generación de residuos per cápita. A continuación, va a definir estos dos eventos con datos anteriores para saber la probabilidad de que ocurra un aumento en la población o un aumento en la generación de residuos per cápita (o ambos).

Tabla 6
Probabilidades de cada evento

Evento	Descripción	Probabilidad
A	Aumento en la población.	0,60
B	Aumento en la generación de residuos per cápita.	0,40
$A \cap B$	Ambos eventos ocurran simultáneamente.	0,30

Nota. Vélez, D., 2024.

La suma de las probabilidades de los eventos A y B por separado es 1 ($0,60 + 0,40$), lo cual es consistente con el concepto de que la suma de todas las probabilidades en un espacio muestral debe ser igual a 1. Es importante notar que la probabilidad de que ambos eventos ocurran simultáneamente ($A \cap B$) es menor que las probabilidades individuales de cada evento. Esto sugiere que los eventos A y B no son independientes, ya que, si lo fueran, la probabilidad de su intersección sería igual al producto de sus probabilidades individuales ($0,60 \times 0,40 = 0,24$).

Sustituimos los valores en la fórmula:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

$$(A \cup B) = 0,60 + 0,40 - 0,30$$

$$P(A \cup B) = 0,70$$

Por lo tanto, la probabilidad de que ocurra un aumento en la población o un aumento en la generación de residuos per cápita (o ambos) es del 70%.

8.5. Independencia, probabilidad condicional y la regla de la multiplicación



- **Independencia de evento:** utilizando el mismo contexto del ejemplo anterior, imagine que el aumento en la población no está directamente relacionado con el aumento en la generación de residuos y viceversa. En este caso, los eventos A y B son independientes.
- **Probabilidad condicional:** $P(B|A)$, probabilidad de que haya un aumento en la generación de residuos dado que ha ocurrido un aumento en la población. Suponga que $P(B|A) = 0,80$, lo que significa que, si hay un aumento en la población, hay un 80 % de probabilidad de que también haya un aumento en la generación de residuos.
- **Regla de la multiplicación:** $P(A \cap B)$, probabilidad de que ocurran ambos eventos, es decir, aumento en la población y aumento en la generación de residuos.

Suponga que $(A) = 0,60$, lo que significa que hay un 60 % de probabilidad de que ocurra un aumento en la población.

Utilizando la regla de la multiplicación $P(A \cap B) = P(A) \times P(B|A)$, puede calcular:

$$P(A \cap B) = 0,60 \times 0,80 = 0,48.$$

Por lo tanto, en este ejemplo, aunque los eventos A y B son independientes, la probabilidad condicional indica que, si hay un aumento en la población, hay una alta probabilidad de que también haya un aumento en la generación de residuos. La regla de la multiplicación le permite calcular la probabilidad de que ambos eventos ocurran simultáneamente.

8.6. Variables aleatorias discretas y sus distribuciones de probabilidad



Imagine un proyecto ambiental que se centra en la conservación de una especie de mariposa en un área protegida. En este caso, la variable aleatoria X pudiera ser el “número de mariposas observadas en una hora”.

Ahora, suponga que ha realizado varias salidas de monitoreo y registrado el número de mariposas observadas en cada hora. Puede construir la distribución de probabilidad discreta que muestra las probabilidades asociadas con cada posible cantidad de mariposas observadas:

- Monitoreo 1: $P(X=0) = 0,10$.
- Monitoreo 2: $P(X=1) = 0,30$.
- Monitoreo 3: $P(X=2) = 0,40$.
- Monitoreo 4: $P(X=3) = 0,15$.
- Monitoreo 5: $P(X=4) = 0,05$.

En este caso, la distribución de probabilidad indica que hay un 10 % de probabilidad de no avistar ninguna mariposa [$P(X=0)$], un 30 % de probabilidad de avistar una mariposa [$P(X=1)$], un 40 % de probabilidad de avistar dos mariposas [$P(X=2)$], un 15 % de probabilidad de avistar tres mariposas [$P(X=3)$], y un 5 % de probabilidad de avistar cuatro mariposas [$P(X=4)$] en una hora de monitoreo. Gráficamente, lo puede representar como se muestra en la figura 12.

Figura 12

Probabilidades de observar mariposas en una hora.



Nota. Vélez, D., 2024.

La figura muestra un histograma que representa la distribución de frecuencias del número de mariposas observadas en una hora. El eje horizontal representa el número de mariposas observadas, que va desde 0 hasta 4. El eje vertical muestra la probabilidad correspondiente para cada valor del número de mariposas.

La probabilidad más alta es observar dos mariposas en una hora y la probabilidad más baja es observar cuatro mariposas en una hora (figura 12). Este ejemplo ilustra cómo las variables aleatorias discretas, como el número de mariposas observadas, pueden modelar eventos en la gestión ambiental y cómo se puede describir la variabilidad de estos eventos mediante una distribución de probabilidad discreta.

Pero, ¿podría calcular la probabilidad de observar al menos una mariposa en una hora? Por supuesto, para calcular la probabilidad de observar al menos una mariposa en una hora, necesitaríamos sumar las probabilidades de observar una o más mariposas en esa hora. Dado que tiene la distribución de probabilidad discreta para el número de mariposas observadas en cada hora, puede simplemente sumar las probabilidades asociadas con observar una, dos, tres o cuatro mariposas.

Entonces, la probabilidad de observar al menos una mariposa en una hora sería:

$$p(X \geq 1) = p(x=1) + p(x=2) + p(x=3) + p(x=4)$$
$$(x \geq 1) = 0,30 + 0,40 + 0,15 + 0,05 = 0,90$$

Por lo tanto, la probabilidad de observar al menos una mariposa en una hora es del 90 %.

8.6.1. La media y desviación estándar para una variable aleatoria discreta

Ahora, con el “número de mariposas observadas en una hora”, calcule la media, la varianza y la desviación estándar de esta variable aleatoria discreta. La fórmula para obtener la media es $\mu = \sum xp(x)$, para la varianza es

$\sigma^2 = \sum (x - \mu)^2 p(x)$ y para la desviación estándar es $\sqrt{\sigma^2}$. Para comprender mejor el proceso de cálculo de la media, varianza y desviación estándar, puede fijarse cada paso en la tabla 7.

Tabla 7

Cálculos para obtener la media, la varianza y desviación estándar de una variable aleatoria discreta.

x	p(x)	xp(x)	$(x - \mu)^2$	$(x - \mu)^2 p(x)$
0	0,10	0,00	3,06	0,306
1	0,30	0,30	0,56	0,169
2	0,40	0,80	0,06	0,025
3	0,15	0,45	1,56	0,234
4	0,05	0,20	5,06	0,253
Total	1,00	$\mu = 1,75$		$\sigma^2 = 0,988$
				$\sigma = 0,994$

Nota. Adaptado de Introducción a la probabilidad y estadística, por W. Mendenhall, 2015, Cengage Learning Editores.

La media indica que se observan 1,75 mariposas por hora, y la desviación estándar mide la dispersión de este número alrededor de la media, siendo aproximadamente 0,994 mariposas (ver tabla 7).

Es momento de aplicar sus conocimientos a través de las actividades que se han planteado a continuación:



Actividades de aprendizaje recomendadas

1. Para fortalecer su comprensión acerca de las variables aleatorias discretas y sus distribuciones de probabilidad, revise el video: "[Función de probabilidad de variable aleatoria discreta](#)".

Como se observó en el ejercicio del video, tiene que encontrar la distribución de probabilidad o función de probabilidad en forma de tabla a partir de la fórmula de una función de probabilidad. También se repasa las condiciones que tiene que cumplir una función para que efectivamente sea una función de probabilidad.

2. Para fortalecer la comprensión de los conceptos de esta unidad, analice y resuelva la siguiente autoevaluación.



Autoevaluación 8

Responda verdadero o falso.

1. La probabilidad es una herramienta importante en estadística, ya que permite cuantificar la incertidumbre y tomar decisiones en situaciones aleatorias.
 - a. Verdadero.
 - b. Falso.

Estimado estudiante, desarrolle los siguientes ejercicios:

2. **Ejercicio 1:** está estudiando la mortalidad de plántulas de una especie de árbol en un vivero. La variable aleatoria X representa el “número de plántulas que mueren por mes”. La siguiente información es la distribución de probabilidad discreta asociada:

$$P(X = 0) = 0,30; P(X = 1) = 0,40; P(X = 2) = 0,20; P(X = 3) = 0,10.$$

Con esta información, responda a las siguientes preguntas:

- a. ¿Cuál es la probabilidad de que ninguna plántula muera en un mes?
- b. ¿Cuál es la probabilidad de que exactamente 2 plántulas mueran en un mes?



c. ¿Cuál es la probabilidad de que al menos una plántula muera en un mes?

3. **Ejercicio 2:** en un estudio de contaminación del agua, la variable aleatoria discreta Y representa el número de eventos de contaminación detectados en un mes. La distribución de probabilidad es la siguiente:

$$P(Y = 0) = 0,20; P(Y = 1) = 0,50; P(Y = 2) = 0,20; P(Y = 3) = 0,10.$$

Responda a las siguientes preguntas:

a. ¿Cuál es la probabilidad de no detectar contaminación en un mes?

b. ¿Cuál es la probabilidad de detectar al menos 2 eventos de contaminación en un mes?

c. ¿Cuál es la probabilidad de detectar exactamente 1 evento de contaminación en un mes?

4. **Ejercicio 3:** en un proyecto de conservación, la variable aleatoria Z representa el número de especies de aves observadas en una hora. La distribución de probabilidad es la siguiente:

$$P(Z = 0) = 0,10; P(Z = 1) = 0,20; P(Z = 2) = 0,40; P(Z = 3) = 0,20; P(Z = 4) = 0,10.$$

Responda a las siguientes preguntas:

a. ¿Cuál es la probabilidad de observar al menos 3 especies de aves en una hora?

b. ¿Cuál es la probabilidad de no observar ninguna especie de ave en una hora?

c. ¿Cuál es la probabilidad de observar exactamente 2 especies de aves en una hora?

[Ir al solucionario](#)



Semana 12

Unidad 9. Algunas distribuciones discretas útiles

Las distribuciones de probabilidad discretas como la binomial y de Poisson desempeñan un papel crucial en la carrera de gestión ambiental al proporcionar herramientas estadísticas para modelar eventos y fenómenos relacionados con la biodiversidad, la conservación y la gestión de recursos naturales. La distribución binomial es esencial para analizar variables con resultados binarios, como la presencia o ausencia de una especie en un área protegida. Por otro lado, la distribución de Poisson es valiosa para modelar eventos que ocurren en una unidad determinada de tiempo o espacio y que se describen con números enteros como la ocurrencia de incendios forestales o la propagación de enfermedades en poblaciones silvestres o en cautiverio. Estas distribuciones, al proporcionar herramientas precisas para la predicción y comprensión de eventos ambientales, le ayudarán a tomar decisiones y diseñar estrategias de conservación y sostenibilidad.

La **distribución binomial** modela experimentos que cumplen con cinco características esenciales (Gavilánez Luna, 2021):

- El experimento consiste en n intentos idénticos.
- Cada intento tiene solo dos resultados posibles (éxito o fracaso).
- La probabilidad de éxito (p) es constante para cada intento.
- Los intentos son independientes entre sí.
- Interesa el número total de éxitos en los n intentos.

Un experimento que cumple estas condiciones se denomina "experimento binomial".

Para el cálculo de probabilidades binomiales, la probabilidad de obtener exactamente k éxitos en n intentos se calcula mediante la fórmula:

$$P(X = k) = C(n,k) \times p^k \times (1-p)^{n-k}$$

Donde:

n es el número total de intentos.

k es el número de éxitos.

p es la probabilidad de éxito en cada intento.

$C(n,k)$ es el coeficiente binomial que representa las combinaciones posibles.

La **distribución de Poisson** modela el número de eventos que ocurren en un intervalo fijo (tiempo, distancia, área o volumen), cuando estos eventos son:

- Independientes entre sí.
- Ocurren con una tasa media constante.
- No pueden ocurrir exactamente al mismo tiempo o lugar.

La probabilidad de observar exactamente k eventos cuando la media es λ se calcula mediante:

$$P(X = k) = (e^{-\lambda} \times \lambda^k) / k!$$

Donde:

k es el número de eventos.

λ (lambda) es el número promedio de eventos por intervalo.

e es la base del logaritmo natural (≈ 2.71828).

Para comprender los conceptos de esta unidad, puede emplear las siguientes estrategias de estudio:

- Asegúrese de comprender los conceptos básicos relacionados con las distribuciones de probabilidad discretas.
- Dedique tiempo a entender las características específicas que definen:
 - Un experimento binomial.

- Un proceso de Poisson.
- Para la distribución binomial, enfóquese en:
 - Las cinco condiciones necesarias.
 - La diferencia entre éxito y fracaso en el contexto del experimento.
 - El cálculo de probabilidades usando la fórmula binomial.
- Para la distribución de Poisson, preste atención a:
 - Las condiciones para su aplicación.
 - El concepto de tasa media de ocurrencia.
 - El cálculo de probabilidades usando la fórmula de Poisson.
- Practique la identificación de situaciones donde cada distribución es apropiada en el contexto de la gestión ambiental.

9.1. La distribución binomial de probabilidad

Considere la variable aleatoria X como el “número de aves observadas en un período de tiempo determinado”, con la cual deberá evaluar la probabilidad de observar exactamente dos aves durante un período específico.

Va a evaluar si este experimento cumple con las cinco características de un experimento binomial:

- El experimento consiste en n intentos idénticos: si cumple, el experimento involucra un número fijo de intentos, que son los diferentes períodos de tiempo durante los cuales se realizará el monitoreo.
- La probabilidad de éxito en un solo intento es igual a p y es el mismo de un intento tras otro: si cumple, asumimos que la probabilidad de observar un ave durante un período específico es constante.
- Los intentos son independientes: si cumple, suponga que la observación de aves en un período no está relacionada con la observación en otros períodos.



- Estamos interesados en X , el número de éxitos observado durante los n intentos, para $x = 0, 1, 2\dots, n$: si cumple, estamos interesados en cuántas aves se observan (número de éxitos) durante los diferentes períodos de observación.



Sin embargo, la segunda característica, que indica que solo debe haber dos resultados posibles (éxito o fracaso) en cada intento, no se cumple completamente en este caso. Puede tener varios resultados posibles, como ninguna ave, una, dos, tres, etc. Esto hace que el experimento no sea completamente binomial.



9.1.1. La distribución binomial de probabilidad



Imagine un proyecto en el cual se está llevando a cabo un estudio de supervivencia de plántulas de guayacán (especie forestal) en un área de reforestación. Se establece un experimento donde se seleccionan al azar 15 plántulas recién plantadas y se monitorea su supervivencia después de un año. Dado que la variable de interés (la supervivencia) tiene solo dos resultados (supervivencia o no supervivencia), puede aplicar la distribución binomial de probabilidad.



Defina las siguientes condiciones para este experimento:

- **Número de intentos (n):** se seleccionan 15 plántulas, lo que representa el número total de intentos en el experimento.
- **Éxito (p):** defina “éxito” como la supervivencia de una plántula después de un año. Suponga que la probabilidad de que una plántula sobreviva es del 70 %, entonces $p = 0,70$.
- **Fracaso (q):** el fracaso es lo opuesto al éxito. En este caso, sería la probabilidad de que una plántula no sobreviva, por lo que $q = 1 - p = 0,30$.



Con esta información, se puede usar la distribución binomial para calcular probabilidades específicas. Por ejemplo, si quiere encontrar la probabilidad de que exactamente 10 plántulas sobrevivan, $P(x = 10)$, usaría la fórmula:

$$P(x=k) = C_k^n p^k q^{n-k}, \text{ donde:}$$

- $P(x=k)$ es la probabilidad de que ocurran k éxitos (plántulas que sobreviven) en n ensayos.
- n es el número total de ensayos o eventos, en este caso, 15 plántulas.
- k es el número de éxitos que se desea calcular (en este caso 10 plántulas).
- p es la probabilidad de éxito en un solo ensayo.
- q ($1-p$) es la probabilidad de fracaso en un solo ensayo.

Reemplazando los valores en la fórmula, tiene:

$$P(x=10) = \left(\frac{15}{10}\right) \times (0,70)^{10} \times (0,30)^5$$

Donde $\left(\frac{15}{10}\right)$ es el coeficiente binomial, que representa el número de formas de seleccionar 10 éxitos de 15 intentos y se calcula como:

$$\left(\frac{15}{10}\right) = \frac{15!}{10! \times (15-10)!}$$

Realizando los cálculos:

$$\left(\frac{15}{10}\right) = \frac{15!}{10! \times 5!} = \frac{15 \times 14 \times 13 \times 12 \times 11}{5 \times 4 \times 3 \times 1} = 3003$$

Sustituya este valor en la fórmula de la distribución binomial:

$$P(x=10) = 3003 \times (0,70)^{10} \times (0,30)^5$$

Calcule las potencias y multiplique:

$$P(x=10) = 3003 \times 0,028 \times 0,002$$

Finalmente, multiplicará estos valores para obtener la probabilidad:

$$P(x=10) \approx 0,211 \quad (\text{El símbolo "≈" significa "aproximadamente igual a".})$$



Por lo tanto, la probabilidad de que exactamente 10 plántulas sobrevivan después de un año es aproximadamente 0,211 o un 21.1 %. A largo plazo, una baja tasa de supervivencia podría no ser suficiente para alcanzar los objetivos de reforestación. Sería necesario revisar y ajustar los métodos de plantación para mejorar la eficiencia de la reforestación.

9.1.2. Media y desviación estándar para la variable aleatoria binomial

Para calcular la media (μ) y la desviación estándar (σ) de la distribución binomial que modela la supervivencia de 15 plántulas de guayacán con una probabilidad de éxito (supervivencia) $p= 0,80$, aplicará las fórmulas:

$$\mu = n \times p, \text{ y};$$

$$\sigma = \sqrt{n \times p \times q}; \text{ recordemos que } q = 1 - p$$

Donde:

- n es el número de ensayos o experimentos (15 en este caso).
- p es la probabilidad de éxito en un solo ensayo (0,80 en este caso).

Sustituya los valores:

$$\mu = 15 \times 0,80 = 12$$

$$\sigma = \sqrt{15 \times 0,80 \times 0,20} = \sqrt{2,4} \approx 1,549$$

Por lo tanto, la media (μ) de la distribución es 12 y la desviación estándar (σ) es aproximadamente 1,549. Estos valores indican que, en promedio, se esperaría que sobreviviesen alrededor de 12 plántulas de cada 15 sembradas, con una variabilidad de aproximadamente 1,549 en la distribución. ¿Qué pasaría si la probabilidad de supervivencia fuera menor?, ¿el promedio de supervivencia de 12 plántulas se mantendría igual? Le propongo repetir este ejercicio, pero ahora con $p = 0,30$.

9.2. La distribución de probabilidad de Poisson



Para comprender la distribución de Poisson con un ejemplo, imagine que estamos estudiando la frecuencia de incendios forestales en una reserva natural. Suponga que, en promedio, ocurren tres incendios por mes en esta área.

La media (μ) en la distribución de Poisson es el número promedio de eventos en el intervalo dado. En este caso, $\mu = 3$ incendios por mes. La desviación estándar (σ) en la distribución de Poisson es la raíz cuadrada de la media. En este caso, $\sigma = \sqrt{3}$.

Ahora su interés es que no haya incendios. Va a averiguar, ¿cuál es la probabilidad de que no haya incendios en esta reserva durante un mes?

Para resolver este problema usando la distribución de Poisson, puede utilizar la fórmula:

$$p(x = k) = \frac{\mu^k \times e^{-\mu}}{k!}$$

Donde:

- $P(x = k)$ es la probabilidad de que haya exactamente k incendios.
- μ es la media de la distribución de Poisson (en este caso, el número promedio de incendios por mes).
- e es la base del logaritmo natural (aproximadamente 2,71828).
- k es el número de incendios que está considerando.

En este caso, quiere encontrar la probabilidad de que no haya incendios ($k=0$) en un mes, y la media μ es 3 incendios por mes. Sustituya estos valores en la fórmula:

$$P(x = 0) = \frac{3^0 \times e^{-3}}{0!}$$

$$P(x = 0) = \frac{1 \times e^{-3}}{1} = e^{-3} \approx 0,049$$

Por lo tanto, la probabilidad de que no haya incendios en la reserva natural durante un mes es aproximadamente 0,0498 o un 4.98 %. Los incendios en áreas naturales cuyos ecosistemas no están adaptados al fuego pueden tener impactos significativos en el medioambiente, incluida la pérdida de biodiversidad, daño a los *hábitats* naturales y cambios en la composición del suelo. La baja probabilidad de no tener incendios resalta la importancia de implementar medidas de prevención y gestión para proteger la flora, fauna y el ecosistema en general.

Para calcular probabilidades binomiales y de Poisson puede utilizar Excel a través de las siguientes funciones:

Para distribución binomial:

- $\text{BINOM}(\text{número_éxitos}, \text{ensayos}, \text{probabilidad_éxito}, \text{acumulado})$. Ejemplo: $=\text{DISTR.BINOM}(5,10,0.5,\text{FALSO})$ calcula la probabilidad de obtener exactamente 5 éxitos en 10 ensayos con probabilidad de éxito 0.5.

Para distribución de Poisson:

- $\text{DIST}(x, \text{media}, \text{acumulado})$. Ejemplo: $=\text{POISSON.DIST}(3,2,\text{FALSO})$ calcula la probabilidad de obtener exactamente 3 eventos cuando el promedio es 2.

Donde:

- El argumento "acumulado" es Falso para calcular probabilidades exactas.
- El argumento "acumulado" es Verdadero para calcular probabilidades acumuladas.

Para más detalles sobre el uso de estas funciones, puede consultar la ayuda de Excel o recursos en línea sobre análisis estadístico.

Es momento de aplicar sus conocimientos a través de la actividad que se ha planteado a continuación:



Actividad de aprendizaje recomendada



Estimado estudiante, para fortalecer la comprensión de los conceptos de esta unidad, analice y resuelva la siguiente autoevaluación:



Autoevaluación 9

Responda verdadero o falso:

1. La distribución binomial es útil para analizar eventos con resultados binarios en la gestión ambiental.
 - a. Verdadero.
 - b. Falso.

2. La distribución de Poisson se utiliza exclusivamente para modelar eventos en el tiempo, no en el espacio continuo.
 - a. Verdadero.
 - b. Falso.

3. Las distribuciones de probabilidad discretas no tienen relevancia en la gestión ambiental.
 - a. Verdadero.
 - b. Falso.

4. La distribución de Poisson es adecuada para modelar eventos que no pueden contarse en números enteros.
 - a. Verdadero.
 - b. Falso.

5. Las distribuciones binomiales y de Poisson son irrelevantes para la toma de decisiones en gestión ambiental.
 - a. Verdadero.

- b. Falso.
6. Las distribuciones de probabilidad discretas son útiles solo para la predicción y no para la comprensión de eventos ambientales.
- a. Verdadero.
- b. Falso.

Realice los siguientes ejercicios:

7. **Ejercicio 1:** retomando el ejemplo de la sección 9.1.1, en donde se seleccionó al azar 15 plántulas recién plantadas y se monitoreó su supervivencia después de un año. ¿Cuál será la probabilidad de que en lugar de 10 plántulas sobrevivieran todas las plántulas, si la probabilidad de supervivencia es del 70 %?
8. **Ejercicio 2:** continuando con el ejercicio anterior (pregunta 7) ¿Cuál será la probabilidad de que sobreviva solamente una plántula de guayacán?
9. **Ejercicio 3:** retomando el ejemplo de la sección 9.2, en donde en promedio ocurren tres incendios por mes en una reserva natural. ¿Cuál será la probabilidad de que haya seis incendios en esta reserva?
10. **Ejercicio 4:** suponga que el promedio de lluvias en una ciudad es cuatro eventos de lluvia por semana. Calcule la probabilidad de que ocurran exactamente un evento de lluvia la próxima semana.

[Ir al solucionario](#)



¡Felicitaciones por completar esta unidad de estudio! ¡Su comprensión y habilidades en estadística están creciendo cada día más!



Semana 13

Unidad 10. La distribución normal de probabilidad

Como las distribuciones de probabilidad binomial y de Poisson, la distribución normal de probabilidad es una herramienta para modelar y comprender fenómenos naturales y procesos ambientales. Su utilidad radica en su capacidad para describir una gama amplia de variables ambientales, proporcionando una representación matemática de cómo los datos cuantitativos distribuyen los valores alrededor de un promedio.

En la gestión ambiental, la distribución normal se utiliza para analizar datos tan diversos como niveles de contaminantes, variaciones climáticas, biomasa de organismos, dinámica de nutrientes, impactos antrópicos y muchos indicadores biológicos. Permite evaluar la probabilidad de eventos extremos o inusuales, así como comprender la variabilidad natural en sistemas ecológicos. Además, facilita la toma de decisiones al proporcionar un marco estadístico para la interpretación de datos y la implementación de estrategias de conservación.

Una **variable aleatoria continua** puede tomar cualquier valor dentro de un rango específico. Las dos distribuciones continuas fundamentales son (Gavilánez Luna, 2021):

- Variable aleatoria uniforme:
 - Todos los valores dentro del rango tienen igual probabilidad de ocurrir.
 - Se define por sus valores mínimo y máximo.
 - Útil para modelar procesos donde cualquier valor en un rango es igualmente probable.

Ejemplo: la concentración permitida de arsénico en agua potable debe estar entre 0 y 10 partes por billón (ppb). Si asumimos una distribución uniforme, cualquier valor en este rango tiene la misma probabilidad de ocurrir.

- Variable aleatoria exponencial:

- Modela el tiempo entre eventos.
- Caracterizada por una tasa de ocurrencia constante.
- Útil para análisis de supervivencia y procesos de llegada.

Ejemplo: El tiempo entre eventos de precipitación ácida en una región industrial puede seguir una distribución exponencial, donde la tasa promedio es de 3 eventos por mes.

La **distribución normal**, también conocida como distribución gaussiana, es una de las distribuciones de probabilidad más importantes en estadística (Douglas, 2003). Sus características principales son:

- Forma de campana simétrica.
- Definida por dos parámetros: media (μ) y desviación estándar (σ).
- La media, mediana y moda coinciden.
- Aproximadamente el 68 % de los datos están dentro de $\pm 1\sigma$ de la media.
- Aproximadamente el 95 % de los datos están dentro de $\pm 2\sigma$ de la media.
- Aproximadamente el 99.7 % de los datos están dentro de $\pm 3\sigma$ de la media.

Un ejemplo de distribución normal son las mediciones de pH en un lago que durante un año suelen seguir una distribución normal con media $\mu = 7,0$ y desviación estándar $\sigma = 0,2$.

Para calcular probabilidades en una distribución normal es necesario:

- Estandarizar los valores usando la fórmula: $Z = (X - \mu)/\sigma$.
- Utilizar tablas de la distribución normal estándar o software estadístico.
- Interpretar los resultados en el contexto del problema.

Un ejemplo de distribución normal de probabilidad es un proceso industrial, el cual debe mantener una temperatura entre 180 °C y 200 °C, y el proceso sigue una distribución normal con $\mu = 190$ °C y $\sigma = 5$ °C, podemos calcular la probabilidad de que la temperatura esté dentro de los límites aceptables.

La **variable aleatoria normal estándar** (Z) es una transformación de cualquier variable aleatoria normal (X) que facilita el cálculo de probabilidades. Esta transformación se obtiene restando la media y dividiendo por la desviación estándar: $Z = (X - \mu)/\sigma$

La distribución normal estándar tiene las siguientes características:

- Media $\mu = 0$: el centro de la distribución está en cero.
- Desviación estándar $\sigma = 1$: la dispersión está estandarizada.
- Se denota como $Z \sim N(0, 1)$: indica que Z sigue una distribución normal con media 0 y desviación estándar 1.

Esta estandarización es útil porque:

- Permite comparar valores de diferentes distribuciones normales al convertirlos a una escala común.
- Simplifica el uso de tablas de probabilidad al necesitar solo una tabla estándar.
- Facilita la interpretación de qué tan lejos está un valor de la media en términos de desviaciones estándar.

Excel proporciona dos funciones principales para el cálculo de probabilidades normales:

1. NORM.N(x, media, desv_estándar, acumulado)

- x : valor para el cual se calcula la probabilidad.
- media: promedio de la distribución.
- desv_estándar: desviación estándar de la distribución.
- acumulado: VERDADERO para $P(X \leq x)$, FALSO para $P(X = x)$.

2. NORM.ESTAND(z)

- z: valor estandarizado.
- Calcula probabilidades directamente para la distribución normal estándar.
- Equivalente a DISTR.NORM.N(z, 0, 1, acumulado).

Por ejemplo, suponga que la concentración de un contaminante sigue una distribución normal con media $\mu = 50$ mg/L y desviación estándar $\sigma = 5$ mg/L. Para calcular la probabilidad de que la concentración sea menor o igual a 55 mg/L:

1. Estandarizamos: $z = (55 - 50)/5 = 1$

2. En Excel:

- Usando DISTR.NORM.N: =DISTR.NORM.N(55, 50, 5, VERDADERO).
- Usando DISTR.NORM.ESTAND: =DISTR.NORM.ESTAND(1).

Ambas funciones darán el mismo resultado: aproximadamente 0,8413, indicando que hay un 84,13% de probabilidad de que la concentración sea menor o igual a 55 mg/L.



Para comprender los conceptos de esta unidad, revise los ejemplos que ilustren cómo se aplican estas distribuciones en situaciones reales. Resuélvalos usted mismo.

10.1. Distribuciones de probabilidad para variables aleatorias continuas

Tanto la variable aleatoria uniforme como la variable aleatoria exponencial son tipos de variables aleatorias continuas. La característica principal de una variable aleatoria continua es que puede tomar un número infinito de valores dentro de su rango.

Para comprender la *variable aleatoria uniforme*, suponga que hemos establecido un límite permitido de concentración de mercurio en el agua, digamos, entre 2 y 5 microgramos por litro ($\mu\text{g/L}$), de acuerdo con las normativas ambientales. La característica principal de la variable aleatoria uniforme en este contexto es que cualquier valor dentro de ese rango tiene la misma probabilidad de ocurrir. Es decir, la probabilidad de que la concentración de mercurio esté exactamente en 2,5 $\mu\text{g/L}$ sería la misma que la probabilidad de que esté en 4 $\mu\text{g/L}$. Esta uniformidad en la probabilidad reflejaría la idea de que no hay razón para esperar concentraciones de mercurio más probables en una ubicación específica a lo largo del río en nuestro escenario hipotético.

En cambio, una *variable aleatoria exponencial*, es como un “cronómetro natural” que nos ayuda a medir el tiempo que tarda en ocurrir un evento, como la germinación de semillas. Imagine que está llevando a cabo un proyecto de reforestación y está interesado en entender cuánto tiempo tarda una semilla en germinar. Después de analizar los datos, descubre que la tasa de germinación promedio es 0,1 germinaciones por día. Con esta información, puede usar la distribución exponencial para modelar el tiempo (en días) que tomará para que una semilla germene. Por ejemplo, podría calcular la probabilidad de que una semilla germene en los primeros 5 días. Esta información puede ser valiosa para la planificación, permitiéndole optimizar las condiciones de germinación y anticipar cuándo se esperaría que germinasen las semillas de cierta especie de interés.

10.2. La distribución normal de probabilidad

Imagine que está estudiando la contaminación por metales pesados en un área afectada por la minería. En este contexto, puede utilizar las distribuciones de probabilidad para modelar la concentración de plomo (Pb) en el agua del río afectado, un indicador crítico de la contaminación.

En este sentido, quiere entender la probabilidad de que la concentración de plomo en una descarga específica sea menor que el límite máximo permisible de 0,2 mg/L, de acuerdo con las regulaciones ambientales.

Si asume que la concentración de plomo sigue una distribución normal con una media (μ) de 0,15 mg/L y una desviación estándar (σ) de 0,05 mg/L, puede utilizar la función DISTR.NORM.N de Excel para calcular la probabilidad de que la concentración de plomo en una descarga específica sea menor que el límite máximo permisible de 0,2 mg/L. Siga los siguientes pasos:

1. Abra Excel y cree una nueva hoja de cálculo.
2. En una celda, ingrese la media (μ) de la distribución normal, que es 0,15 mg/L.
3. En otra celda, ingrese la desviación estándar (σ) de la distribución normal, que es 0,05 mg/L.
4. En una celda separada, ingrese el límite máximo permisible de concentración de plomo, que es 0,2 mg/L.
5. Utilice la función de Excel DISTR.NORM.N para calcular la probabilidad asociada a ese valor utilizando la distribución normal. La sintaxis de esta función es: DISTR.NORM.N (x, media, desv_estándar, acumulado), donde:
 - x es el valor para el cual desea calcular la probabilidad (en este caso, el límite máximo permisible).
 - media es la media de la distribución normal.
 - desv_estándar es la desviación estándar de la distribución normal, y
 - acumulado es un valor lógico que indica si desea la probabilidad acumulada o no acumulada. En este caso, establezca este argumento en verdadero o 1 para obtener la probabilidad acumulada.
6. Ejecute (con Enter) la función DISTR.NORM.N con los valores correspondientes y obtendrá la probabilidad deseada.

La probabilidad de que la concentración de plomo sea menor que 0,2 mg/L es aproximadamente 0,84. Esto significa que la mayoría de las descargas tendrán concentraciones de plomo por debajo del límite máximo permisible, lo que indicaría un cumplimiento favorable de las regulaciones ambientales en cuanto a la concentración de plomo en las descargas.

10.3. Áreas tabuladas de la distribución normal de probabilidad

Las áreas tabuladas de la distribución normal de probabilidad son herramientas prácticas que simplifican los cálculos relacionados con la distribución normal, lo que facilita el análisis de datos y la toma de decisiones en una variedad de contextos estadísticos.

10.3.1. La variable aleatoria normal estándar

La variable aleatoria normal estándar es una variable aleatoria continua que sigue una distribución normal (en forma de montículo) con una media (promedio) de cero y una desviación estándar de uno. Se utiliza principalmente en estadística para estandarizar otras distribuciones normales. Esto significa que cualquier variable aleatoria normal con cualquier media y desviación estándar podría convertirse en una variable aleatoria normal estándar mediante el proceso de estandarización. Esto es útil para comparar diferentes conjuntos de datos que pueden tener diferentes medias y desviaciones estándar.

10.3.2. Cálculo de probabilidades para una variable aleatoria normal general

Para comprender la *variable aleatoria normal* estándar y el *cálculo de probabilidades para una variable aleatoria normal general*, va a emplear un ejemplo. Suponga que un laboratorio ambiental está realizando pruebas de calidad del aire en una ciudad para monitorear los niveles de un contaminante específico, como el plomo (Pb). Se sabe que la concentración de plomo sigue una distribución normal con una media de $30 \text{ } \mu\text{g}/\text{m}^3$ y una desviación estándar de $2 \text{ } \mu\text{g}/\text{m}^3$.

Para calcular la probabilidad de que una muestra de aire tenga una concentración de plomo superior a $40 \text{ } \mu\text{g}/\text{m}^3$, primero necesita estandarizar la variable utilizando la variable aleatoria normal estándar. La fórmula para estandarizar una variable es:

$$z = \frac{x-\mu}{\sigma}$$

Donde:

- z es la variable aleatoria normal estándar.
- x es el valor de la variable cuya probabilidad se quiere calcular.
- μ es la media de la distribución original.
- σ es la desviación estándar de la distribución original.

En este caso, quiere calcular $P(X>50)$, así que primero estandariza $50 \text{ }\mu\text{g/ m}^3$:

$$z = \frac{40-30}{2} = 5$$

Ahora que ha estandarizado la variable, puede consultar la tabla de distribución normal estándar para encontrar la probabilidad asociada con $z = 5$. Según la tabla, esta probabilidad es: $P(z>5) = 0$. Usando una tabla de distribución normal estándar, encontrará que la probabilidad de que z sea mayor que 5 es prácticamente cero. Esto significa que la probabilidad de que una muestra de aire tenga una concentración de plomo superior a $40 \text{ }\mu\text{g/m}^3$ es muy cercana a cero.

Este ejemplo ilustra cómo se puede utilizar la variable aleatoria normal estándar y el cálculo de probabilidades para analizar datos relacionados con la gestión ambiental, como la calidad del aire de una ciudad.

Continuemos con el aprendizaje mediante su participación en la actividad que se describe a continuación:



Actividad de aprendizaje recomendada

Para fortalecer la comprensión de los conceptos de esta unidad, analice y resuelva la siguiente autoevaluación.



Autoevaluación 10

Responda verdadero o falso.

1. La distribución normal de probabilidad es una herramienta útil para modelar y comprender fenómenos naturales y procesos ambientales.
 - a. Verdadero.
 - b. Falso.
2. La distribución normal no proporciona un marco estadístico para la interpretación de datos en la gestión ambiental.
 - a. Verdadero.
 - b. Falso.
3. Tanto la variable aleatoria uniforme como la variable aleatoria exponencial son tipos de variables aleatorias continuas.
 - a. Verdadero.
 - b. Falso.
4. Las áreas tabuladas de la distribución normal de probabilidad son útiles solo en contextos estadísticos simples y no en análisis de datos más complejos.
 - a. Verdadero.
 - b. Falso.
5. La variable aleatoria normal estándar sigue una distribución normal con una media de cero y una desviación estándar de uno.
 - a. Verdadero.
 - b. Falso.



Desarrolle los siguientes ejercicios y responda:

6. Ejercicio 1. Bacterias en agua potable

Suponga que la cantidad de un tipo particular de bacteria, en muestras de 1 mililitro (ml) de agua potable, tienden a estar normalmente distribuidas en forma aproximada con una media de 85 y una desviación estándar de 9. ¿Cuál es la probabilidad de que una muestra determinada de 1 ml contenga más de 100 bacterias



7. Ejercicio 2. Cambio climático

Se sabe que la temperatura media mensual en un ecosistema seco está normalmente distribuida con una media de 25 °C y una desviación estándar de 3 °C. ¿Cuál es la probabilidad de que la temperatura en un mes dado esté por encima de 30 °C?



8. Ejercicio 3. Polinizadores

En un estudio sobre la diversidad de los insectos polinizadores en un área afectada por el cambio climático, se encontró que el número promedio de especies de abejas por sitio de muestreo sigue una distribución normal con una media de 25 especies y una desviación estándar de 5 especies.



- ¿Cuál es la probabilidad de encontrar más de 30 especies de abejas en un sitio de muestreo?
- ¿Cuál es la probabilidad de encontrar entre 20 y 30 especies de abejas en un sitio de muestreo?
- ¿Qué porcentaje de los sitios de muestreo tendrán 15 o menos especies de abejas?



[Ir al solucionario](#)



¡Felicitaciones por completar esta unidad de estudio! ¡Está cada vez más cerca de dominar el arte de la estadística y su impacto en el mundo que nos rodea!



Unidad 11. Distribuciones muestrales

Recuerde que, en la estadística, la población se refiere al grupo completo que se quiere estudiar, mientras que la muestra se refiere a un subconjunto representativo de esa población. La distribución muestral de una estadística es la distribución de probabilidad para los posibles valores de la estadística, que resulta cuando muestras se sacan repetida y aleatoriamente de la población. Son herramientas estadísticas que nos permiten entender la variabilidad de las muestras y cómo se relacionan con la población general. Al estudiar las distribuciones muestrales, los profesionales en gestión ambiental pueden tomar decisiones basadas en la comprensión de la incertidumbre inherente a los datos ambientales. Esto les permite diseñar políticas, estrategias y acciones de conservación más efectivas para preservar y proteger los ecosistemas.

La **distribución muestral** es fundamental en estadística, ya que describe cómo varían los estadísticos calculados a partir de diferentes muestras de una misma población. Cuando diferentes investigadores toman muestras de una población, aunque utilicen el mismo tamaño de muestra y metodología, obtendrán valores ligeramente diferentes para los estadísticos calculados (como la media o la proporción). La distribución de estos valores forma lo que conocemos como distribución muestral (Melo et al., 2007).

Los **planes muestrales** y los **diseños experimentales** son herramientas metodológicas complementarias en la investigación cuantitativa. Un plan muestral establece la estrategia para seleccionar unidades representativas de una población, considerando aspectos como la definición precisa de la población objetivo, el método apropiado de selección, el tamaño adecuado de la muestra y los procedimientos estandarizados de medición (Domínguez & Tostado, 2016).



Por ejemplo, en un estudio de calidad del aire urbano, un plan muestral podría definir la población como el aire de toda la ciudad, emplear un muestreo estratificado por zonas (industrial, residencial y comercial), establecer 30 puntos de muestreo distribuidos proporcionalmente entre las zonas, y especificar mediciones estandarizadas de PM_{2.5} usando equipos calibrados según normas internacionales.

Por otro lado, un **diseño experimental** define la estructura y organización de un experimento que permita establecer relaciones causales válidas. Los elementos esenciales de un diseño experimental incluyen la identificación clara de variables dependientes e independientes, el control riguroso de variables extrañas que podrían afectar los resultados, la aleatorización en la asignación de tratamientos y la replicación adecuada para garantizar la validez estadística (Gutiérrez Pulido & Vara Salazar, 2012).

Para ilustrar, en un estudio de biorremediación de suelos contaminados, el diseño experimental podría establecer como variable independiente el tipo de tratamiento (con o sin bacterias específicas), como variable dependiente el nivel de contaminantes en el suelo, incluir parcelas de control sin tratamiento, asignar aleatoriamente los tratamientos a diferentes parcelas y realizar cinco repeticiones por tratamiento para asegurar la confiabilidad de los resultados.

La **distribución muestral** de un estadístico describe el comportamiento de ese estadístico cuando se calcula repetidamente en diferentes muestras del mismo tamaño extraídas de una población. Esta distribución tiene tres propiedades fundamentales que son su forma, su centro y su dispersión. Estas propiedades nos permiten entender cómo varía el estadístico de muestra a muestra y hacer inferencias sobre la población (Douglas, 2003).

En el contexto ambiental, por ejemplo, piense en un programa de monitoreo de la calidad del agua en un río donde se seleccionan aleatoriamente 10 puntos de muestreo cada mes para medir el oxígeno disuelto. Aunque cada conjunto mensual de muestras proporcionará una media diferente, la distribución de



estas medias mensuales (la distribución muestral) nos ayuda a entender la variabilidad natural en nuestras estimaciones y a hacer inferencias más precisas sobre la verdadera calidad del agua en el río.

El **teorema del límite central** es uno de los principios fundamentales en estadística que establece que cuando el tamaño de la muestra es suficientemente grande, la distribución muestral de la media se aproxima a una distribución normal, independientemente de la forma de la distribución de la población original. Este teorema es particularmente importante porque proporciona la base teórica para realizar inferencias estadísticas sobre poblaciones usando información muestral (Gavilánez Luna, 2021).

La aplicación de este teorema es especialmente relevante en estudios ambientales donde las poblaciones originales pueden tener distribuciones complejas o desconocidas. Por ejemplo, en el monitoreo de contaminantes atmosféricos, las mediciones individuales pueden no seguir una distribución normal, pero si tomamos muestras lo suficientemente grandes (generalmente $n \geq 30$), las medias de estas muestras se distribuirán aproximadamente de forma normal, permitiendo el uso de métodos estadísticos basados en la normalidad.

La **distribución muestral de la media** describe cómo varían las medias calculadas a partir de todas las posibles muestras de un mismo tamaño extraídas de una población. Esta distribución tiene propiedades importantes que la caracterizan: su media es igual a la media poblacional, y su desviación estándar (conocida como error estándar) es igual a la desviación estándar poblacional dividida por la raíz cuadrada del tamaño de la muestra (Gutiérrez Pulido & Vara Salazar, 2012).

El **error estándar** es una medida crucial de la precisión de nuestras estimaciones muestrales. Se define como la desviación estándar de la distribución muestral y nos indica cuánta variación podemos esperar en nuestras estimaciones de una muestra a otra. Un error estándar más pequeño indica estimaciones más precisas, y podemos reducirlo aumentando el tamaño de la muestra (Douglas, 2003).

En estudios ambientales, el error estándar es particularmente útil para determinar la confiabilidad de nuestras mediciones. Por ejemplo, al monitorear los niveles de contaminación en un cuerpo de agua, un error estándar pequeño nos daría más confianza en que nuestra media muestral está cerca del verdadero valor poblacional. Si el error estándar es grande, podríamos necesitar aumentar nuestro tamaño de muestra o mejorar nuestros métodos de medición para obtener estimaciones más precisas.

La importancia práctica de estos conceptos se evidencia cuando necesitamos tomar decisiones basadas en datos muestrales. Por ejemplo, si estamos evaluando si un ecosistema cumple con ciertos estándares de calidad ambiental, necesitamos considerar no solo el valor promedio de nuestras mediciones, sino también su error estándar para establecer intervalos de confianza apropiados y tomar decisiones informadas sobre la gestión ambiental (Melo et al., 2007).

Para revisar estos contenidos le recomiendo realizar lo siguiente:

- Divida cada tema en secciones más pequeñas y establezca un plan de estudio detallado.
- Mientras estudia, tome notas de los conceptos clave, definiciones y ejemplos importantes. Luego, sintetice esta información en resúmenes breves que pueda revisar fácilmente.
- Resuelva problemas y ejercicios relacionados con cada tema.
- Encuentre ejemplos prácticos o casos de estudio relacionados con la gestión ambiental para aplicar los conceptos de estadística en un contexto relevante.
- Preste especial atención a conceptos como el teorema del límite central, la distribución muestral de la media muestral y el error estándar, ya que son fundamentales en estadística.
- Si encuentra dificultades con algún tema o concepto, no dude en buscar ayuda con su profesor.

Siguiendo esta estrategia, podrá estudiar de manera efectiva estos temas y mejorar su comprensión de ellos.

11.1. Planes muestrales y diseños experimentales



Imagine que está trabajando en un proyecto de monitoreo de la calidad del agua en un río cercano a una planta industrial. La planta ha estado operando durante varios años y existe preocupación sobre el impacto de sus actividades en el ecosistema acuático. Para evaluar adecuadamente esta situación, es fundamental diseñar un plan muestral y un experimento bien estructurado.

En este caso, un plan muestral detallado podría incluir la selección de sitios de muestreo representativos a lo largo del río, considerando diferentes condiciones ambientales y posibles fuentes de contaminación. Además, se establecerían criterios claros para la recolección de muestras de agua, como la profundidad, el caudal y la frecuencia de muestreo.

Por otro lado, el diseño experimental implicaría establecer grupos de control y tratamiento para comparar la calidad del agua, aguas arriba y aguas abajo de la planta industrial. Se podrían implementar diferentes estrategias de muestreo, como *muestreo aleatorio* o *estratificado*, para garantizar la representatividad de los datos recopilados.

Un diseño experimental bien planificado permitiría identificar si hay cambios significativos en la calidad del agua asociados con la actividad de la planta industrial y proporcionaría evidencia sólida para la toma de decisiones en términos de regulación ambiental y gestión de riesgos.



Imagine que la planta industrial mencionada anteriormente vierte sus efluentes en un río cercano, y existe preocupación por la posible contaminación del agua con pesticidas. Para evaluar el impacto de la planta en la calidad del agua, se lleva a cabo un estudio utilizando un diseño muestral y experimental adecuado.

Se seleccionan dos puntos de muestreo: uno aguas arriba de la planta industrial y otro aguas abajo. En cada punto, se recogen 25 muestras de agua en diferentes momentos a lo largo de un período de tiempo determinado.

Después de analizar las muestras, se calculan los promedios y las desviaciones estándar de la concentración de pesticidas en cada conjunto de muestras. Los resultados muestran que el promedio de concentración de pesticidas aguas arriba es menor que el promedio aguas abajo. Además, la desviación estándar de las muestras aguas abajo es mucho mayor que la desviación estándar de las muestras aguas arriba, lo que sugiere una mayor variabilidad en la concentración de pesticidas en esta ubicación.

Estos hallazgos son importantes para la gestión ambiental porque indican que la planta industrial puede estar contribuyendo a niveles más altos y menos predecibles de contaminación por pesticidas en el agua aguas abajo. Esto resalta la importancia de implementar medidas de control y monitoreo adecuadas en la planta para reducir su impacto ambiental y proteger la calidad del agua en la región. Además, subraya la relevancia de un diseño muestral y experimental bien planificado para obtener resultados confiables y significativos en la evaluación de la calidad ambiental.

Los planes muestrales y diseños experimentales son fundamentales en la gestión ambiental porque garantizan la recopilación de datos confiables y representativos, lo que facilita la evaluación precisa de los impactos ambientales y la implementación de medidas correctivas adecuadas para proteger y preservar los ecosistemas.

11.2. Estadística y distribuciones muestrales

Imagine que trabaja en un proyecto de restauración de un ecosistema forestal degradado. Para evaluar la eficacia de diferentes técnicas de restauración, decide recolectar datos sobre el crecimiento (de acuerdo a la altura) de los árboles en varias parcelas dentro del área de restauración.

Al analizar estos datos, se enfrenta al desafío de entender la variabilidad en las alturas de los árboles y cómo se distribuyen estas alturas en toda la población de árboles.

Aquí es donde entra en juego la estadística y las distribuciones muestrales. Utilice técnicas estadísticas para resumir y analizar los datos recolectados, como calcular la media, la mediana y la desviación estándar de las alturas de los árboles en sus parcelas de muestreo. Estas medidas descriptivas le ayudan a comprender la distribución de alturas en su muestra.

Además, puede utilizar métodos de inferencia estadística para hacer inferencias sobre la población de árboles en todo el área de restauración. Por ejemplo, podría calcular intervalos de confianza para la altura promedio de los árboles con base en su muestra, lo que le proporcionaría una estimación de la altura promedio de todos los árboles en el área de restauración y la incertidumbre asociada a esta estimación.

En este contexto, la estadística y las distribuciones muestrales le permiten obtener información valiosa sobre la variabilidad y la distribución de características importantes en el ecosistema forestal restaurado. Esta comprensión es fundamental para tomar decisiones sobre las estrategias de restauración y evaluar el éxito del proyecto en términos de recuperación de la estructura y función del ecosistema.

11.3. El teorema del límite central

El *teorema del límite central* es un teorema fundamental para el análisis de muestras grandes. Este teorema establece que, independientemente de la forma de la distribución de la población original, la distribución de las medias muestrales, se aproxima a una distribución normal a medida que el tamaño de la muestra aumenta.



Para comprender el teorema del límite central, imagine que usted está trabajando en un proyecto de monitoreo de la calidad del agua en un río para evaluar la contaminación por nutrientes. Para ello, toma muestras de agua en diferentes puntos a lo largo del río y analiza la concentración de nitratos en cada muestra.

Al recolectar una gran cantidad de muestras y calcular la concentración promedio de nitratos en cada una, se da cuenta de que estas concentraciones muestrales siguen una distribución aproximadamente normal (en forma de montículo). Sin embargo, debido a la variabilidad natural en la concentración de nitratos y otros factores ambientales, las concentraciones individuales pueden variar considerablemente.

En el contexto de su estudio sobre la calidad del agua del río, el teorema del límite central le permite inferir sobre la concentración promedio de nitratos en el río a partir de sus muestras. Puede calcular intervalos de confianza para la concentración promedio de nitratos utilizando la distribución normal de las medias muestrales, lo que le proporciona una estimación precisa y una medida de la incertidumbre asociada.

El *teorema del límite central* es fundamental en la gestión ambiental porque le permite hacer inferencias confiables sobre las características de la población, como la calidad del agua del río, a partir de datos muestrales, lo que facilita la toma de decisiones para la conservación y protección del medioambiente.

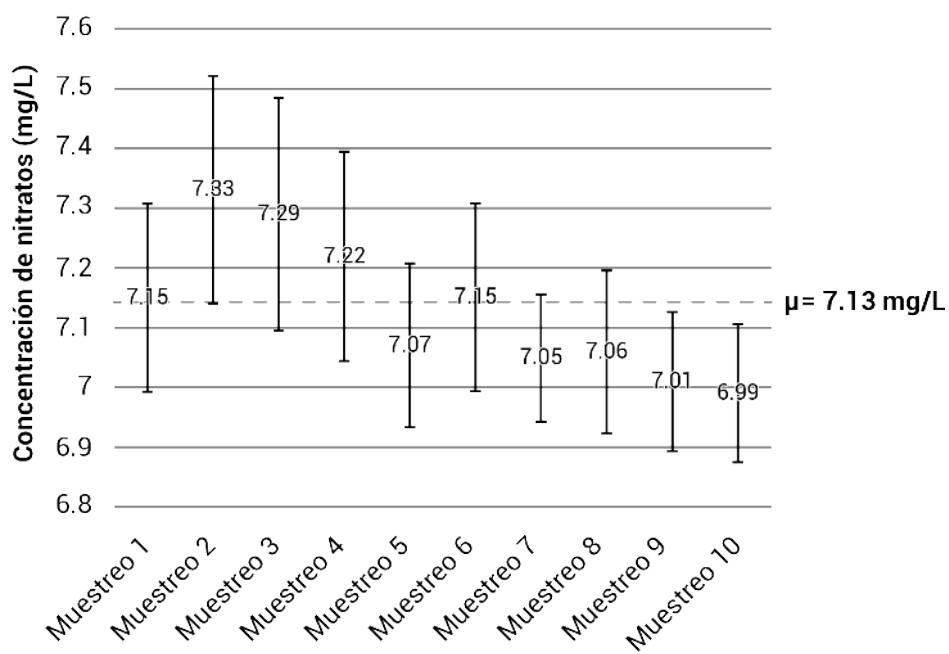
11.4. La distribución muestral de la media muestral

La *distribución muestral de la media muestral* es un concepto fundamental en estadística que describe cómo se distribuyen las medias muestrales de todas las posibles muestras de una población. Esto es importante porque nos permite comprender la variabilidad de la media muestral y proporciona información sobre la precisión de nuestras estimaciones poblacionales basadas en muestras.

Retomando el ejemplo anterior, suponga que recolecta muestras de agua a lo largo de un río para medir la concentración de nitratos en cada muestra. Estime la media poblacional (μ) a partir de las medias () de las diferentes muestras y luego grafique como se muestra en la figura 13. En la figura se observa que la media poblacional (7,13 mg/L de concentración de nitratos) coincide con la mayoría de las distribuciones muestrales de cada una de las muestras, excepto las muestras 9 y 10. Esto sugiere una variabilidad inusual en la concentración de nitratos en esas ubicaciones específicas del río, indicando valores bajos de fuentes locales de contaminación o condiciones ambientales particulares que afectan la concentración de nitratos en esas áreas.

Figura 13

Media poblacional de 10 muestras de agua para medir la concentración de nitratos. Cada muestra tiene una media muestral independiente.



Nota. Vélez, D., 2024.

La figura representa la concentración de nitratos (en mg/L) a lo largo del tiempo, desde el muestreo 1 hasta el muestreo 10. La línea discontinua representa la media general de la concentración de nitratos (en mg/L) de todos los muestreos. Esta línea sirve como punto de referencia para comparar las concentraciones individuales de cada muestreo.

11.4.1. Error estándar

El error estándar es una medida de la dispersión de las medias muestrales alrededor de la verdadera media poblacional. Cuanto menor sea el error estándar, mayor será la precisión de nuestra estimación de la media poblacional. Como puede ver en la figura 13, el tamaño de las barras representa el error estándar. En esta figura, observe que hay mayor error estándar y, por tanto, mayor variación en la muestra 3 y menor variación en la muestra 7.

Es momento de aplicar sus conocimientos a través de la actividad que se ha planteado a continuación:



Actividad de aprendizaje recomendada

Para fortalecer la comprensión de los conceptos de esta unidad, analice y resuelva la siguiente autoevaluación.



Autoevaluación 11

Responda verdadero o falso.

1. En estadística, la población se refiere al grupo completo que se quiere estudiar, mientras que la muestra se refiere a un subconjunto representativo de esa población.
 - a. Verdadero.
 - b. Falso.

2. La estadística y las distribuciones muestrales no proporcionan información relevante sobre la variabilidad y la distribución de características importantes en el ecosistema forestal restaurado.



- a. Verdadero.
- b. Falso.

3. El teorema del límite central es relevante solo para muestras pequeñas.



- a. Verdadero.
- b. Falso.



Desarrolle los siguientes ejercicios:



4. **Ejercicio 1:** se llevó a cabo un estudio sobre la preferencia de actividades recreativas en áreas verdes de una ciudad. Se seleccionó una muestra aleatoria al entrevistar a cada décimo visitante que ingresaba a un parque urbano durante un fin de semana soleado.



- ¿Esta muestra tendrá las características de una muestra aleatoria seleccionada de los visitantes? Explique.

5. **Ejercicio 2:** un equipo de investigadores desea evaluar la exposición a la contaminación del aire en niños que asisten a una escuela ubicada al centro de la ciudad. Deciden llevar a cabo un estudio piloto para evaluar la calidad del aire y sus posibles efectos en la salud de los niños. Seleccionan a 50 niños de entre 8 y 12 años que cumplen con los criterios de elegibilidad del estudio, como residir en la ciudad, asistir regularmente a la escuela y tener consentimiento de sus padres para participar en el estudio. Durante nueve meses, los investigadores monitorean la calidad del aire en áreas cercanas a la escuela y recolectan datos sobre la frecuencia de síntomas como tos, dificultad para respirar o sibilancias.

- a. Describa el plan muestral empleado para seleccionar participantes para el estudio.

- b. ¿Qué mecanismo de probabilidad se utilizó para seleccionar esta muestra de 50 niños de entre 8 a 12 años?
 - c. ¿Se pueden hacer inferencias válidas usando los resultados de este estudio?, ¿por qué?
 - d. Si usted tuviera que idear un plan muestral alternativo, ¿qué cambiaría?
6. **Ejercicio 3:** un equipo de científicos realiza mediciones repetidas de la temperatura del agua en diferentes puntos a lo largo del río utilizando termómetros digitales de alta precisión. A pesar de que el experimento se lleva a cabo de manera idéntica en cada punto, es posible que las mediciones de temperatura obtenidas en repeticiones distintas no sean exactamente iguales. Estas diferencias en las mediciones pueden influir en la interpretación de los resultados del estudio.
- a. Haga una lista de algunas variables de este experimento que podrían causar algunos pequeños cambios en la medición de la respuesta final.
 - b. Si se desea verificar que el error de medición es pequeño, se puede repetir el experimento y tomar el promedio muestral de todas las mediciones. Para reducir la cantidad de variabilidad en el promedio de mediciones, ¿debe usarse un número de repeticiones grande o pequeño? Explique.

[Ir al solucionario](#)



¡Siga adelante! Comprender las distribuciones muestrales es la clave para desbloquear un mundo de posibilidades en el análisis estadístico.

Resultado de aprendizaje 6:

Escribe reportes de investigación siguiendo las normas técnicas y científicas.

El resultado de aprendizaje le permitirá comunicar de manera efectiva los hallazgos de sus investigaciones en el campo ambiental. Los reportes de investigación son vitales para compartir resultados, conclusiones y recomendaciones derivadas de investigaciones ambientales, lo que contribuye a la generación de conocimiento en el campo, facilita la toma de decisiones y promueve acciones efectivas para la conservación y gestión sostenible del medioambiente.

Contenidos, recursos y actividades de aprendizaje recomendadas

Recuerde revisar de manera paralela los contenidos con las actividades de aprendizaje recomendadas y actividades de aprendizaje evaluadas.



Semana 15

Unidad 12. El reporte de investigación

El reporte de investigación permite al estudiante comunicar de manera efectiva los resultados de sus estudios observacionales o experimentales. Estos informes proporcionan un medio para compartir hallazgos, conclusiones y recomendaciones derivadas de la investigación, lo que contribuye al avance del conocimiento en el campo ambiental y facilita la toma de decisiones para la conservación y gestión sostenible del ambiente. Además, los reportes de investigación son herramientas clave para la difusión de información científica entre la comunidad académica, los responsables de políticas ambientales y el público en general, promoviendo la conciencia y la acción en beneficio del ambiente.

El **reporte de investigación** es un documento formal que comunica de manera sistemática y objetiva los hallazgos, métodos y conclusiones de un estudio científico (Munch Galindo, 2009). En el campo de la gestión ambiental, estos reportes son fundamentales para compartir conocimientos sobre problemas ambientales, metodologías de evaluación y estrategias de conservación.

La estructura del reporte de investigación típicamente incluye las siguientes secciones:

- Introducción: la introducción establece el contexto y relevancia del estudio, presenta el problema de investigación y los objetivos. Por ejemplo, en un estudio sobre contaminación atmosférica, la introducción describiría la problemática actual de la calidad del aire en zonas urbanas, su impacto en la salud pública y los objetivos específicos del estudio (Hernández-Sampieri, 2018).
- Marco Teórico: esta sección presenta la base teórica y revisión de literatura que fundamenta el estudio. Se analizan investigaciones previas, teorías relevantes y conceptos clave relacionados con el tema de investigación. En un estudio ambiental, podría incluir teorías sobre dispersión de contaminantes, efectos en la salud y normativas ambientales aplicables (Hernández-Sampieri, 2018).
- Metodología: la metodología describe detalladamente los procedimientos utilizados para recopilar y analizar datos. Esta sección debe ser lo suficientemente clara y detallada para que otros investigadores puedan replicar el estudio. En gestión ambiental, incluye aspectos como el diseño de muestreo, métodos de análisis, equipos utilizados y procedimientos de control de calidad. Por ejemplo, en un estudio de calidad del agua, se especificarían los puntos de muestreo, frecuencia de medición, parámetros analizados y métodos analíticos empleados (Hernández-Sampieri, 2018).
- Resultados: los resultados presentan los hallazgos del estudio de manera objetiva, utilizando texto, tablas, gráficos y figuras según sea apropiado. Esta sección debe organizarse lógicamente siguiendo los objetivos de investigación, sin incluir interpretaciones. En estudios ambientales, es común presentar datos cuantitativos como concentraciones de contaminantes, índices de biodiversidad o tendencias temporales,

acompañados de análisis estadísticos relevantes (Hernández-Sampieri, 2018).

- Discusión: la discusión interpreta los resultados en el contexto del conocimiento existente. Aquí se explica el significado de los hallazgos, se comparan con estudios previos y se discuten las implicaciones prácticas. Por ejemplo, en un estudio sobre especies amenazadas, la discusión podría analizar cómo los resultados contribuyen a la comprensión de factores que afectan la supervivencia de la especie y sus implicaciones para la conservación (Hernández-Sampieri, 2018).
- Conclusiones y recomendaciones: esta sección sintetiza los principales hallazgos y su significado, respondiendo a los objetivos planteados inicialmente. Las recomendaciones deben ser específicas y basadas en la evidencia presentada. En gestión ambiental, podrían incluir sugerencias para políticas de conservación, medidas de mitigación o áreas que requieren investigación adicional.
- Referencias: las referencias documentan todas las fuentes citadas en el reporte, siguiendo un formato consistente (como APA, Vancouver o Chicago). Deben ser actualizadas y relevantes para el tema de estudio.
- Anexos: los anexos incluyen material suplementario que, aunque importante, interrumpiría el flujo principal del texto. Pueden incluir datos crudos, cálculos detallados, fotografías, mapas o documentación adicional relevante para el estudio (Domínguez & Tostado, 2016).

Al redactar un reporte de investigación ambiental, es fundamental:

- Mantener un estilo claro, conciso y objetivo.
- Usar terminología técnica apropiada pero accesible.
- Presentar datos de manera organizada y visualmente efectiva.
- Mantener la coherencia entre las diferentes secciones.
- Seguir las normas de formato y citación requeridas.
- Considerar las necesidades y conocimientos de la audiencia objetivo.

La efectividad de un reporte de investigación no solo depende de la calidad del estudio realizado, sino también de cómo se comunican los hallazgos de manera clara y convincente para diferentes audiencias, desde científicos hasta tomadores de decisiones en gestión ambiental.

Para revisar estos contenidos le recomiendo realizar lo siguiente:

- Dedique tiempo a comprender por qué cada sección es importante en un reporte de investigación. Por ejemplo, la introducción establece el contexto y la motivación del estudio, mientras que la discusión interpreta los resultados y los relaciona con la literatura existente.
- Observe cómo las secciones están interconectadas entre sí. Por ejemplo, la introducción presenta la pregunta de investigación y los objetivos, que luego son abordados en la sección del método y discutidos en la sección de resultados y discusión.

12.1. Reporte de resultados del proceso de investigación

El reporte de investigación es un documento que requiere una cuidadosa planificación y estructuración para cumplir con su propósito. Antes de su elaboración, es fundamental definir claramente a los usuarios a los que se destina, ya que esto determinará el formato, la naturaleza y la extensión del informe. Los reportes pueden presentarse en contextos académicos o no académicos, lo que también influye en su diseño y contenido. En el caso de un reporte académico, la estructura común incluye portada, índices, resumen, introducción, revisión de la literatura, métodos, análisis y resultados, discusión, referencias y apéndices.

A continuación, verá un ejemplo de la estructura de un reporte de investigación a través de su índice. Este ejemplo es una revisión de estudios acerca del grado de investigación relacionado con los factores de cambio global en el matorral interandino de Ecuador.

Revisión de los factores de cambio global en el matorral interandino de Ecuador



Introducción



Revisión de la literatura

- Estado de conocimiento del matorral seco interandino.
- Principales factores de cambio global.
- Los Objetivos de Desarrollo Sostenible (ODS).
- Importancia del estudio de los factores de cambio global en el matorral interandino en Ecuador.
- Objetivos del estudio.



Métodos



- Área de estudio.
- Recolección de datos.
- Criterios de exclusión en la búsqueda.
- Factores de cambio global.
- Clasificación de los estudios.
- Análisis de datos.



Resultados

- Distribución de investigaciones relacionadas con factores de cambio global.
- Distribución de investigaciones relacionadas con los Objetivos de Desarrollo Sostenible (ODS).
- Número de investigaciones relacionadas con los factores de cambio global.
- Número de investigaciones relacionadas con los Objetivos de Desarrollo Sostenible (ODS).
- Número de investigaciones por taxón y relacionadas con los factores de cambio global.
- Número de investigaciones por taxón relacionadas con los Objetivos de Desarrollo Sostenible (ODS).



Discusión

- Factores de cambio global más estudiados.

- Factores de cambio global menos estudiados.
- Objetivos de Desarrollo Sostenible más estudiados.
- Objetivos de Desarrollo Sostenible menos estudiados.
- Taxones más estudiados.
- Taxones menos estudiados.
- Conclusiones.
- Recomendaciones.

Referencias

De esta manera podemos observar un ejemplo de un índice de un reporte de investigación.

Continuemos con el aprendizaje mediante su participación en las actividades que se describen a continuación.



Actividades de aprendizaje recomendadas

1. Localice un artículo de una revista científica y analice las secciones del artículo. ¿Cuál sería su índice? (Los artículos normalmente no tienen, pero puede deducirlo).

Retroalimentación:

Una posible respuesta podría ser del siguiente estilo: para realizar el índice para este ejercicio me basé en el estudio denominado “Estimación de la población de perros callejeros en el distrito de Los Olivos, Lima, Perú”. Los autores son: Yaliní Ochoa A., Néstor Falcón P., José Zuazo R., y Beder Guevara P.

• Introducción

- Definición de perro callejero.
- Problemas asociados con los perros callejeros.
- Importancia de controlar la población de perros callejeros.
- Objetivo del estudio.

• **Antecedentes y revisión de la literatura**

- Situación de los perros callejeros en América Latina.
- Legislación y políticas relacionadas con la tenencia de animales.
- Estudios previos sobre poblaciones de perros callejeros.



• **Materiales y métodos**

- Área de estudio: Distrito de Los Olivos, Lima.
- Diseño del estudio: observacional, transversal y descriptivo.
- Selección de la muestra y muestreo probabilístico.
- Procedimientos de recolección de datos.
- Estimación de la población de perros callejeros.
- Análisis estadístico de los resultados.



• **Resultados**

- Número estimado de perros callejeros.
- Distribución de perros por zonas.
- Características de los perros callejeros (sexo, tamaño, estado de bienestar).



• **Discusión**

- Interpretación de los resultados.
- Implicaciones para la salud pública y el bienestar animal.
- Comparación con estudios previos y consideraciones metodológicas.
- Conclusiones.
- Recomendaciones.



• **Referencias**

2. Para fortalecer la comprensión de los conceptos de esta unidad, analice y resuelva la siguiente autoevaluación.





Autoevaluación 12

1. El reporte de investigación es una herramienta efectiva para comunicar los resultados de estudios observacionales o experimentales.
 - a. Verdadero.
 - b. Falso.

2. Los informes de investigación no contribuyen al avance del conocimiento en el campo ambiental.
 - a. Verdadero.
 - b. Falso.

3. Los informes de investigación no facilitan la toma de decisiones para la conservación y gestión sostenible del ambiente.
 - a. Verdadero.
 - b. Falso.

4. Los reportes de investigación no son herramientas importantes para la difusión de información científica entre la comunidad académica y los responsables de políticas ambientales.
 - a. Verdadero.
 - b. Falso.

5. Los reportes de investigación no tienen impacto en la promoción de la conciencia y la acción en beneficio del ambiente entre el público en general.
 - a. Verdadero.
 - b. Falso.

6. El reporte de investigación no requiere planificación ni estructuración para cumplir con su propósito.



- a. Verdadero.
b. Falso.
7. Es fundamental definir claramente a los usuarios a los que se destina el reporte antes de su elaboración.
- a. Verdadero.
b. Falso.
8. Los reportes de investigación solo se presentan en contextos académicos, lo que no influye en su diseño y contenido.
- a. Verdadero.
b. Falso.
9. La estructura común de un reporte académico incluye solo cinco secciones: introducción, métodos, resultados, discusión y conclusiones.
- a. Verdadero.
b. Falso.
10. La revisión de la literatura incluye una sección sobre la distribución de investigaciones relacionadas con los Objetivos de Desarrollo Sostenible (ODS).
- a. Verdadero.
b. Falso.

[Ir al solucionario](#)



¡Felicitaciones por llegar hasta el final de la revisión y estudio de esta asignatura! ¡Siga adelante con su aprendizaje y nunca subestime el poder de la estadística y la investigación para impulsar su crecimiento y desarrollo!



Resultado de aprendizaje 4 a 6:

- Levanta, analiza e interpreta información de campo.
- Genera investigaciones científicas y aplica sus resultados como aporte a la solución de problemas ambientales.
- Escribe reportes de investigación siguiendo las normas técnicas y científicas.

Contenidos, recursos y actividades de aprendizaje recomendadas



Semana 16

Actividades finales del bimestre

Estimado estudiante, revise la siguiente estrategia de estudio diseñada especialmente para ayudarle a prepararse de manera efectiva para la evaluación del segundo bimestre. A través de un enfoque estructurado y actividades clave, podrá repasar los contenidos estudiados, fortalecer su comprensión y adquirir las habilidades necesarias para destacar en la evaluación. ¡Aproveche al máximo esta oportunidad de aprendizaje y prepárese para alcanzar el éxito académico!

Estrategia de estudio para el segundo bimestre:

1. Organice un calendario de estudio:

- Divida los contenidos de las unidades 7 a 12 en bloques manejables.
- Asigne tiempo específico para estudiar cada unidad y realizar las actividades de aprendizaje recomendadas.

2. Repase los conceptos fundamentales:

- Revise detenidamente los apuntes y materiales de cada unidad.
- Haga énfasis en los conceptos clave, como probabilidad, distribuciones de probabilidad, variables aleatorias y distribuciones muestrales.
- Cree resúmenes o esquemas que sinteticen la información esencial.

3. Realice las actividades de aprendizaje recomendadas:

- Complete las actividades propuestas para cada unidad, ya que refuerzan la comprensión de los temas.
- Aproveche estas actividades para aplicar los conceptos aprendidos y desarrollar habilidades prácticas.



4. Resuelva problemas y ejercicios:

- Practique el cálculo de probabilidades, la interpretación de distribuciones de probabilidad y la aplicación de teoremas importantes, como el teorema del límite central.
- Utilice los ejemplos proporcionados en clase y busque problemas adicionales para fortalecer sus habilidades.



5. Complete las autoevaluaciones:

- Realice las autoevaluaciones de las unidades 7 a 12, para medir su comprensión de los temas.
- Identifique las áreas en las que necesita reforzar su conocimiento y dedique tiempo adicional a estudiarlas.



6. Participe activamente en la tutoría:

- Involúcrese en las discusiones y actividades propuestas por el docente.
- Plantee preguntas para aclarar dudas y compartir sus ideas con los compañeros.
- Aproveche la retroalimentación del docente para mejorar su comprensión.



7. Aplique los conceptos a situaciones reales:

- Busque ejemplos de aplicación de los conceptos aprendidos en contextos reales, como estudios de investigación o casos prácticos.
- Analice cómo se utilizan las distribuciones de probabilidad y las técnicas de muestreo en esos escenarios.

8. Repase y consolide su aprendizaje:

- Dedique tiempo regularmente a repasar los contenidos estudiados.
- Utilice tarjetas de estudio (*flashcards*) para memorizar definiciones, fórmulas y conceptos clave.
- Realice pruebas de práctica para evaluar su progreso y prepararse para la evaluación.



Recuerde ser constante y disciplinado en su estudio, y no dude en buscar ayuda cuando la necesite. Con dedicación y esfuerzo, estará bien preparado para demostrar su dominio de los contenidos en la evaluación del segundo bimestre. ¡Éxito en su proceso de aprendizaje!





4. Autoevaluaciones

Autoevaluación 1

Pregunta	Respuesta	Retroalimentación
1 Ejercicio 1 Caso 1		Este caso tiene un enfoque cualitativo, ya que, la investigación se centra en comprender las percepciones, experiencias y opiniones de la comunidad local en relación con el proyecto de instalación de un parque eólico en una zona rural.
2 Ejercicio 1 Caso 2		Este caso tiene un enfoque mixto, ya que evaluar los indicadores del ecosistema fluvial implica la medición y cuantificación de variables específicas relacionadas con la calidad del agua y la salud del ecosistema, pero además implica evaluar las percepciones de la comunidad sobre la calidad del agua, sus usos y prácticas relacionadas.
3 Ejercicio 2 Determinar los problemas de contaminación de las mecánicas.		El objetivo es muy general y poco específico. No especifica qué tipo de problemas, qué tipos de contaminación, o si se trata de algún sitio en particular.
4 Ejercicio 2 Disminuir los niveles de contaminación provocada por los automotores.		Términos generales, poco específicos. Objetivo dirigido a una consecuencia, entregable, producto o impacto de la investigación. No especifica el tipo de contaminación, qué automotor. ¿Es el objetivo o el resultado de la investigación?
5 Ejercicio 2 ¿Cuál es el impacto ambiental en la Amazonía durante el último año de determinada empresa minera?		Pregunta que no implica una investigación completa, sino la obtención de un dato o cierta información. Se limita a la búsqueda de un dato específico.



Pregunta	Respuesta	Retroalimentación
6 Ejercicio 2 Elaborar estrategias efectivas de conservación para especies en peligro de extinción.		Objetivo dirigido a una consecuencia, entregable, producto o impacto de la investigación. No es un objetivo de investigación, sino una etapa del proceso investigativo.
7 Ejercicio 2 Aplicar un cuestionario y una prueba estadística para analizar la diferencia entre hombres y mujeres en cuanto a la percepción sobre el cambio climático.		Términos generales, poco específicos. Objetivo dirigido a una etapa del proceso investigativo y no a todo el proceso. ¿En qué contexto?, ¿aplicar un cuestionario y una prueba estadística constituyen un objetivo de investigación?
8 Ejercicio 2 ¿De qué manera se puede educar a los niños de primaria de la ciudad de Guayaquil, Ecuador?		Términos generales, poco específicos. Asimismo, la pregunta no es una interrogante de investigación. No conduce a un estudio concreto. ¿Educar qué o respecto a qué? Impreciso.
9 Ejercicio 3 Transforme y desarrolle la idea en un objetivo de investigación.		Objetivo de investigación: determinar el impacto de las prácticas agrícolas intensivas en los ecosistemas sensibles de la región costera de Ecuador.
10 Ejercicio 3 Transforme y desarrolle la idea en una pregunta de investigación cuantitativa, concisa y concreta.		Pregunta de investigación: ¿cuál es el efecto de las prácticas agrícolas intensivas en la región costera de Ecuador sobre la biodiversidad y los servicios ecosistémicos?

[Ir a la autoevaluación](#)

Autoevaluación 2

Pregunta	Respuesta	Retroalimentación
1	V	El marco teórico sirve como fundamento conceptual para la investigación, proporcionando una base sólida de conocimiento previo sobre el tema en cuestión.
2	F	El marco teórico en gestión ambiental juega un papel crucial al proporcionar una estructura conceptual sólida y orientar tanto la investigación como la toma de decisiones relacionadas con la preservación del medioambiente.
3	F	Si bien la base de datos UTPL es una herramienta útil, también se pueden utilizar otros centros de referencia como Google Académico para detectar fuentes primarias mediante la identificación de palabras clave tanto en español como en inglés.
4	V	La revisión de la literatura es fundamental para identificar teorías, generalizaciones, conclusiones y otros hallazgos relevantes que contribuyen a la construcción del marco teórico.
5	V	La elección adecuada de términos de búsqueda mejora la precisión de los resultados obtenidos, facilitando la identificación de la literatura relevante para el objetivo de investigación.
6	F	La selección de la literatura pertinente requiere un análisis más detallado y crítico de las referencias encontradas, considerando su relevancia, calidad y contribución al marco teórico. No es suficiente basarse únicamente en las primeras 30 referencias.
7	V	La primera etapa del proceso consiste en identificar y recopilar la literatura relevante que aborda el problema de investigación planteado.
8	V	Durante esta etapa, se define la organización y estructura del marco teórico, así como la elaboración de un esquema o índice que guíe su desarrollo.





9

Los estudios más significativos relacionados con la calidad del agua del río Machángara de Quito concluyen lo siguiente:

Teoría 1. Contaminación por desechos orgánicos e inorgánicos: se ha encontrado una alta presencia de desechos orgánicos e inorgánicos en el agua del río Machángara, lo que indica una contaminación significativa. Esto incluye residuos de industrias, aguas residuales sin tratar, pesticidas y otros productos químicos.

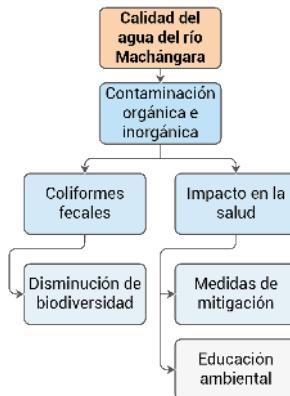
Teoría 2. Niveles elevados de coliformes fecales: los estudios han revelado altos niveles de coliformes fecales en el agua del río, lo que indica una contaminación fecal y la posible presencia de patógenos dañinos para la salud humana.

Teoría 3. Disminución de la biodiversidad acuática: la contaminación del río Machángara ha llevado a una disminución significativa de la biodiversidad acuática, incluyendo la disminución de especies de peces, insectos y otros organismos acuáticos.

Haciendo la comparación con las teorías extraídas de las referencias consultadas, podemos decir que existe una alta contaminación por desechos orgánicos e inorgánicos, niveles elevados de coliformes fecales, disminución de la biodiversidad acuática y riesgos para la salud humana.

10

Basado en las referencias encontradas y utilizando el método de mapeo para construir el marco teórico del objetivo “*Evaluuar la calidad del agua del río Machángara de la ciudad de Quito mediante oxígeno disuelto*” se obtuvo el siguiente mapa conceptual:



Pregunta Respuesta Retroalimentación

Ir a la autoevaluación



Autoevaluación 3

Pregunta	Respuesta	Retroalimentación
1	V	Al definir el alcance de la investigación, se delimita el objetivo y se establecen los límites del estudio, lo que facilita la selección de la metodología más adecuada para alcanzar los objetivos planteados.
2	V	Los estudios exploratorios se centran en investigar fenómenos poco conocidos o en generar una comprensión inicial de un problema, proporcionando información valiosa para futuras investigaciones más detalladas.
3	F	Los estudios descriptivos se centran en observar y describir cómo es y cómo se manifiesta un fenómeno, sin buscar explicaciones causales.
4	F	Los estudios correlacionales se centran en identificar y medir la relación o asociación entre variables, pero no intentan determinar causas y efectos.
5	V	En los estudios explicativos, es esencial comprender los mecanismos subyacentes que explican por qué ocurren ciertos fenómenos y cómo interactúan las variables para producir esos resultados, lo que implica buscar relaciones causales y entender los procesos subyacentes.
6 Ejercicio 1		El objetivo “Evaluar la calidad del agua del río Machángara de la ciudad de Quito mediante oxígeno disuelto” corresponde a un estudio con alcance explicativo porque el objetivo implica medir y describir la calidad del agua utilizando una variable específica, en este caso, el oxígeno disuelto. La descripción se centra en cuantificar la concentración de oxígeno disuelto en el agua del río.
7 Ejercicio 2 a.		<p>Esta pregunta de investigación con alcance descriptivo le permitiría obtener una visión detallada y objetiva de la calidad del aire en el parque industrial durante el periodo señalado, sin buscar explicar las razones detrás de las condiciones observadas.</p> <p>No implica la identificación de relaciones de causa y efecto ni la explicación de por qué ciertas condiciones se observan. El objetivo es proporcionar una descripción detallada.</p>

Pregunta	Respuesta	Retroalimentación
8	Ejercicio 2 b.	Esta pregunta de investigación con alcance correlacional le permitiría analizar si hay una relación entre la densidad vehicular y la concentración de partículas suspendidas en áreas urbanas específicas durante diferentes momentos del día. La investigación correlacional le ayudaría a comprender si estas dos variables están asociadas entre sí. Aunque la pregunta tiene un alcance correlacional, no busca explicar por qué ocurre la relación, solo se busca determinar si existe.
9	Ejercicio 2 c.	Esta pregunta de investigación con alcance exploratorio busca comprender las opiniones, conocimientos y actitudes de los jóvenes con respecto al cambio climático, explorando sus creencias, preocupaciones y posibles acciones de mitigación. Los datos recopilados proporcionarán una descripción de las percepciones de los jóvenes, contribuyendo a una comprensión de la conciencia ambiental en esta población específica.
10	Ejercicio 2 d.	Esta pregunta de investigación, igualmente con alcance explicativo, pretende comprender las razones subyacentes desde diversas perspectivas acerca del abandono de perros para proponer políticas más efectivas en el ámbito de la ecología urbana y el bienestar animal.

[Ir a la autoevaluación](#)

Autoevaluación 4

Pregunta	Respuesta	Retroalimentación
1	Correlacional.	El ejemplo se considera una hipótesis correlacional porque establece una relación entre dos variables.
2	De pronóstico de un dato.	El ejemplo se considera una hipótesis de pronóstico de un dato porque hace una afirmación sobre un evento futuro basado en datos y tendencias actuales.
3	Causal. Dos variables independientes y una dependiente.	El ejemplo se considera una hipótesis causal porque establece una relación de causa y efecto entre dos variables independientes y una variable dependiente.
4	De comparación de grupos.	El ejemplo se considera una hipótesis de comparación de grupos porque establece una comparación entre dos grupos, con respecto a un resultado específico.
5	Correlacional.	El ejemplo se considera una hipótesis correlacional porque establece una relación entre dos variables.
6 Ejercicio 2		Las actividades humanas como la urbanización y la descarga de aguas residuales en los ríos afectan la vida acuática, la biodiversidad y la salud humana.
7 Ejercicio 3 ¿Cuál es la pregunta de investigación de este proyecto?		La pregunta de investigación sería: ¿Cuál es el impacto de las estrategias de sensibilización ambiental en la gestión de residuos sólidos urbanos en la ciudad de Guayaquil?
8 Ejercicio 3 ¿Cuál es la hipótesis que se busca probar por medio de los resultados del experimento?		Las estrategias de sensibilización ambiental en la ciudad de Guayaquil reducen la generación de residuos urbanos.

Pregunta	Respuesta	Retroalimentación
9 Ejercicio 3 ¿Cuál es la variable dependiente y cuál es la variable independiente?		La variable independiente es la implementación de las estrategias de sensibilización ambiental y la variable dependiente es la gestión de residuos sólidos urbanos.
10 Ejercicio 3 ¿Cuántos grupos se incluyen en el experimento?		Se incluyen dos grupos, un grupo experimental y otro grupo de control.
11 Ejercicio 3 ¿Son equivalentes?		Los dos grupos, el grupo experimental y de control son equivalentes porque son hogares que tienen las mismas características socioeconómicas, variando únicamente la manipulación de la variable independiente que es la, implementación de las estrategias de sensibilización ambiental para que las diferencias entre los dos grupos puedan atribuirse a la manipulación de la variable independiente y no a otros factores.
12 Ejercicio 3 ¿Es un diseño experimental o no experimental?		El diseño que se ha elegido es un diseño experimental.
13 Ejercicio 3 ¿Cree que este experimento pueda tener algún efecto importante?		Sí cumple con los tres requisitos de un experimento (1. Manipulación de la variable independiente; 2. Medición de la variable dependiente, y 3. Control sobre la situación experimental) es muy probable que las estrategias de sensibilización ambiental en la gestión de residuos sólidos tengan un efecto importante.

Pregunta	Respuesta	Retroalimentación
14 Ejercicio 4		<p>Para enfoque no experimental, el investigador acudiría a varias zonas de la ciudad de Guayaquil y seleccionaría a una muestra de hogares y encontraría que hay toda una variedad de niveles socioeconómicos de hogares entre ellos (bajo, medio y alto). Luego, a través de encuestas evaluaría el grado de sensibilización ambiental y la gestión de residuos sólidos entre todos los hogares de la muestra y llevaría a cabo sus comparaciones para ver si encuentra tendencias en la gestión de residuos sólidos y así poder establecer la relación entre la sensibilización ambiental y la gestión de residuos sólidos.</p>

[Ir a la autoevaluación](#)



Autoevaluación 5

Pregunta	Respuesta	Retroalimentación
1 Ejercicio 1 Estado de conservación de una reserva natural o parque nacional.		Cualitativa, ya que las categorías del estado de conservación pueden ser: no evaluado, datos insuficientes, preocupación menor, casi amenazado, vulnerable, en peligro, en peligro crítico, extinto en estado silvestre y extinto.
2 Ejercicio 1 Concentración de dióxido de carbono (ppm) en la atmósfera.		Cuantitativa, ya los valores de las concentraciones CO ₂ pueden estar representadas por números.
3 Ejercicio 1 Nivel de conciencia ambiental para medir la percepción y el conocimiento de la población sobre temas ambientales.		Cualitativa, ya que los niveles de conciencia ambiental pueden ser alto, medio y bajo.
4 Ejercicio 1 Cantidad de residuos sólidos urbanos (kg) por día en un hogar.		Cuantitativa, ya que la cantidad de residuos sólidos urbanos (kg) estarán representados por valores numéricos.
5 Ejercicio 1 Consumo de agua potable per cápita medido en litros por persona por día.		Cuantitativa, ya que el consumo de agua potable medido en litros estará representado con números.
6 Ejercicio 2 Temperatura del aire.		Continua, ya que la temperatura del aire puede medirse con precisión decimal, y es una variable que puede tomar cualquier valor dentro de un rango.
7 Ejercicio 2 Número de especies vegetales en un bosque.		Discreta, ya que el número de especies es un valor contable y discreto, y no puede tener fracciones o valores intermedios.
8 Ejercicio 2 Nivel de ruido ambiental en una ciudad.		Continua, ya que el nivel de ruido puede medirse con alta precisión y puede tener valores en un rango continuo.
9 Ejercicio 2 Cantidad de residuos reciclados por mes medido en kg.		Continua, ya que, la cantidad de residuos reciclados puede medirse en kilogramos, puede tener fracciones o valores intermedios.

Pregunta	Respuesta	Retroalimentación
10	Ejercicio 2 Cantidad de incidentes de contaminación ambiental registrados en un año en Ecuador.	Discreta, ya que la cantidad de incidentes de contaminación ambiental contabiliza con números enteros la cantidad de incidentes ocurridos.
11	Ejercicio 3 Con solo mirar la tabla, ¿qué forma piensa usted que tendrá la distribución de datos para concentración de nitratos en el agua de un río (mg/L)?	Solo con mirar la tabla de datos se puede imaginar que se trata de una distribución aproximadamente simétrica, aunque de acuerdo con dos valores altos de concentración de nitratos la distribución estaría sesgada a la derecha.
12	Ejercicio 3 Trace un histograma de frecuencia relativa para describir la distribución. Describa la forma de la distribución y busque resultados atípicos. ¿La gráfica confirma su respuesta al literal a?	De acuerdo con el histograma realizado con una amplitud de intervalo de 0.5 unidades, se generaron cinco intervalos de clase, por tanto, la distribución es simétrica con dos valores atípicos. Se tendría que poner especial cuidados a estos valores altos para verificar si se trata de algún error en la medición o se trata de un aumento en los niveles de concentración de nitratos.

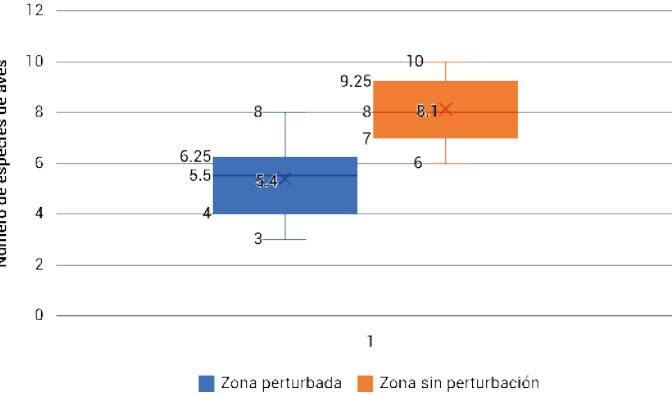
[Ir a la autoevaluación](#)



Autoevaluación 6

Pregunta	Respuesta	Retroalimentación
1	F	La descripción de un conjunto de datos no se limita únicamente al uso de gráficas. Si bien las representaciones gráficas son herramientas poderosas para comprender patrones y tendencias en los datos, el análisis estadístico implica utilizar una combinación entre gráficas y medidas numéricas para obtener una comprensión completa de los datos.
2	Ejercicio 1 ¿Cuáles son la media y desviación estándar para la zona perturbada?	La media y desviación estándar para la zona perturbada es de 5,4 y 1,51, respectivamente.
3	Ejercicio 1 ¿Cuáles son la media y desviación estándar para la zona sin perturbación?	La media y desviación estándar para la zona sin perturbación es de 8,1 y 1,37, respectivamente.
4	Ejercicio 1 Compare los centros y variabilidades de las dos zonas usando los resultados de las partes a y b.	Comparando los centros (con la media) y variabilidades (con la desviación estándar) de las dos zonas, vemos que la media de especies de aves es mayor en la zona sin perturbación, y es menor la variabilidad de acuerdo con su desviación estándar.
5	Ejercicio 1 Encuentre la mediana, los cuartiles superior e inferior y el IQR para cada uno de los dos conjuntos de datos.	Para la zona perturbada la mediana es de 5,5; Q1 = 4; Q3 = 6,25 y IQR = 2,25. Para la zona sin perturbación la media es 8; Q1 = 7; Q3 = 9,25 y IQR = 2,25.



Pregunta	Respuesta Retroalimentación
6 Ejercicio 1 Construya dos gráficas de caja para comparar los dos conjuntos de datos.	 <p>Número de especies de aves</p> <p>1</p> <p>Zona perturbada Zona sin perturbación</p>
7 Ejercicio 1 ¿Qué conclusiones puede extraer sobre la variación del número de especies de aves en las dos zonas, considerando que cada árbol alberga un número diferente de especies de aves?	Se podría decir que los datos son confiables, ya que la variabilidad en los datos es similar en las dos zonas debido a los valores de la desviación estándar y los valores de IQR.
8 Ejercicio 2 Calcule el valor de s usando la aproximación de rango.	Utilizando la aproximación del rango, la desviación estándar para este conjunto de datos es aproximadamente 7.75.
9 Ejercicio 2 Calcule \bar{x} y s y compare con la aproximación de rango de la parte a.	El valor de la media es de 59,2; de la desviación estándar es 10,4. La media y la desviación estándar calculadas utilizando la fórmula estándar proporcionan una medida más precisa de la tendencia central y la dispersión de los datos en comparación con la aproximación del rango. La elección entre estas dos depende de la precisión que se requiera en el análisis de los datos.



Pregunta	Respuesta Retroalimentación
10 Ejercicio 2. Construya una gráfica de caja para los datos y úsela para describir la distribución de datos.	La distribución de los datos no es simétrica y tiene un sesgo a la izquierda, debido a la distancia que hay entre los valores de la mediana (Q2) y el valor más bajo (mín.).

[Ir a la autoevaluación](#)



Autoevaluación 7

Pregunta	Respuesta	Retroalimentación
1	V	Correcto, los datos bivariados involucran dos variables y permiten explorar la relación entre ellas. Esto es fundamental para comprender cómo se relacionan dos fenómenos o características diferentes.
2	V	Las gráficas para variables cualitativas son herramientas visuales utilizadas para representar la distribución de datos que están categorizados o clasificados en diferentes grupos o categorías.
3	F	En las gráficas de dispersión, cada punto en el plano cartesiano representa una observación individual y muestra los valores específicos de ambas variables para esa observación. Esto permite visualizar la relación entre las dos variables para cada observación en el conjunto de datos.
4	F	Las medidas numéricas para datos cuantitativos bivariados, como la correlación o la covarianza, proporcionan información tanto cualitativa como cuantitativa sobre cómo las variables están relacionadas entre sí. Estas medidas cuantifican la fuerza y la dirección de la relación entre las variables, lo que ayuda a entender mejor la naturaleza de la asociación entre ellas.



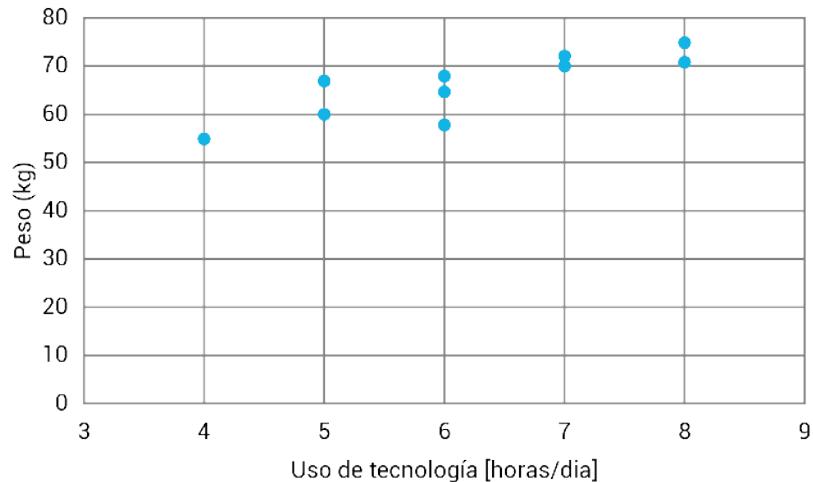
Pregunta

Respuesta Retroalimentación

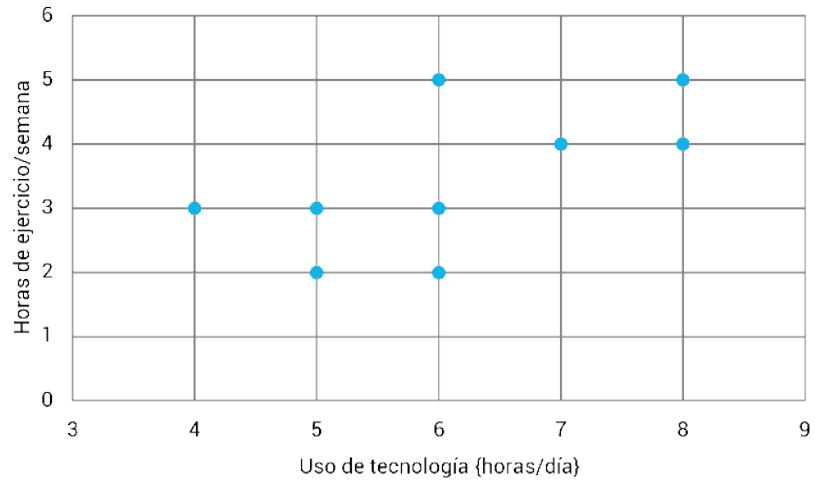


Espera una fuerte relación entre dieta y peso, y entre horas de ejercicio y peso.

Ejercicio
5
1
a.



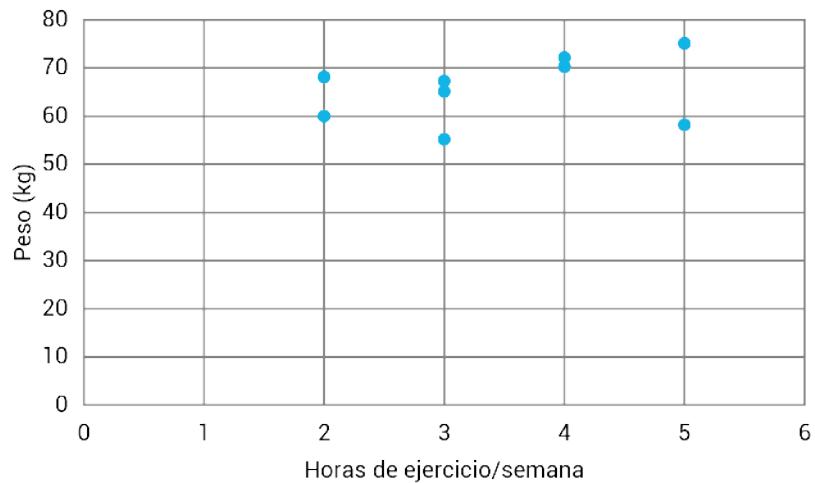
6 Ejercicio
1
b.



Pregunta

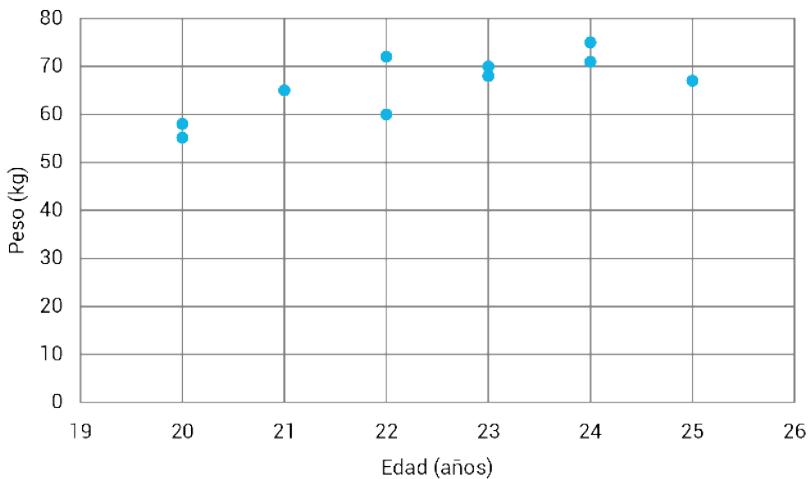
Respuesta Retroalimentación

7 Ejercicio
1
c.





8 Ejercicio
1
d.



A simple vista no se ven conglomerados ni resultados atípicos. El mejor modelo parece la relación entre edad y peso, es decir, a mayor edad, mayor peso.

9 Ejercicio
1
e.

Considere que el par de variables más alineadas son uso de tecnología y peso, las cuales tienen un coeficiente de correlación de 0,84. También, considere que otro par de variables más alineadas son edad y peso, las cuales tienen un coeficiente de correlación de 0,75.

10 Ejercicio
1
f.

Existió una relación lineal positiva entre uso de tecnología y peso, y entre edad y peso de acuerdo con los valores de correlación.

Aunque, la relación entre uso de tecnología y hora de ejercicio, y horas de ejercicio y peso también tuvieron una relación positiva, sus valores de correlación son más bajos ($r = 0,63$ y $r = 0,33$; respectivamente).

[Ir a la autoevaluación](#)

Autoevaluación 8

Pregunta	Respuesta	Retroalimentación
1	V	<p>La probabilidad desempeña un papel crucial en la estadística al proporcionar una forma de cuantificar la incertidumbre en situaciones aleatorias. En la vida cotidiana y en diversos campos como la ciencia, la ingeniería, la economía y la medicina, nos enfrentamos a eventos inciertos cuyos resultados no podemos predecir con certeza. La probabilidad nos permite evaluar la posibilidad de que ocurran diferentes resultados y asignar números que representen esas posibilidades.</p>
2		<p>Ejercicio 1 ¿Cuál es la probabilidad de que ninguna plántula muera en un mes?</p> <p>$P(X=0) = 0,30.$</p>
3		<p>Ejercicio 1 ¿Cuál es la probabilidad de que exactamente 2 plántulas mueran en un mes?</p> <p>$P(X=2) = 0,20.$</p>
4		<p>Ejercicio 1 ¿Cuál es la probabilidad de que al menos una plántula muera en un mes?</p> <p>$P(X \geq 1) = 1 - P(X=0) = 1 - 0,3 = 0,70.$</p>
5		<p>Ejercicio 2 ¿Cuál es la probabilidad de no detectar contaminación en un mes?</p> <p>$P(Y=0) = 0,20.$</p>
6		<p>Ejercicio 2 ¿Cuál es la probabilidad de detectar al menos 2 eventos de contaminación en un mes?</p> <p>$(Y \geq 2) = P(Y=2) + P(Y=3) = 0,20 + 0,10 = 0,30.$</p>

Pregunta	Respuesta	Retroalimentación
7 Ejercicio 2 ¿Cuál es la probabilidad de detectar exactamente 1 evento de contaminación en un mes?	$P(Y=1) = 0,50.$	
8 Ejercicio 3 ¿Cuál es la probabilidad de observar al menos 3 especies de aves en una hora?	$P(Z \geq 3) = P(Z=3) + P(Z=4) = 0,20 + 0,10 = 0,30.$	
9 Ejercicio 3 ¿Cuál es la probabilidad de no observar ninguna especie de ave en una hora?	$P(Z=0) = 0,10.$	
10 Ejercicio 3 ¿Cuál es la probabilidad de observar exactamente 2 especies de aves en una hora?	$P(Z=2) = 0,40.$	

[Ir a la autoevaluación](#)

Autoevaluación 9

Pregunta	Respuesta	Retroalimentación
1	V	La distribución binomial es esencial para analizar variables con resultados binarios, como la presencia o ausencia de una especie en un área protegida.
2	F	La distribución de Poisson se utiliza para modelar eventos que ocurren en una unidad determinada de tiempo o espacio, lo que incluye eventos espaciales y temporales.
3	F	Las distribuciones de probabilidad discretas, como la binomial y la de Poisson, desempeñan un papel crucial en la gestión ambiental al proporcionar herramientas estadísticas para modelar eventos y fenómenos relacionados con la biodiversidad y la conservación.
4	F	La distribución de Poisson es valiosa para modelar eventos que se describen con números enteros, como la ocurrencia de incendios forestales o la propagación de enfermedades en poblaciones silvestres o en cautiverio.
5	F	Estas distribuciones son cruciales para la toma de decisiones en gestión ambiental al proporcionar herramientas precisas para la predicción y comprensión de eventos ambientales.
6	F	Además de ayudar en la predicción, estas distribuciones también son valiosas para comprender mejores eventos ambientales y tomar decisiones informadas.
7 Ejercicio 1:		La probabilidad de que todas las plántulas sobrevivan después de un año es aproximadamente 0,00016 o un 0,016 %.
8 Ejercicio 2:		La probabilidad de que sobreviva solamente una plántula de guayacán es aproximadamente 0,0000000871 o un 0,00000871 %.
9 Ejercicio 3:		La probabilidad de que haya seis incendios en la reserva natural durante un mes es aproximadamente 0,0000701 o un 0,00701 %.
10 Ejercicio 4:		La probabilidad de que ocurra exactamente 1 evento de lluvia en una semana en esta ciudad es 0,073264 o un 7,32 %.

[Ir a la autoevaluación](#)

Autoevaluación 10

Pregunta	Respuesta	Retroalimentación
1	V	La distribución normal es ampliamente utilizada en la gestión ambiental debido a su capacidad para describir una variedad de variables ambientales y proporcionar una representación matemática de la distribución de datos alrededor de un promedio.
2	F	La distribución normal facilita la toma de decisiones en la gestión ambiental al proporcionar un marco estadístico para la interpretación de datos y la implementación de estrategias de conservación.
3	V	Ambas son variables aleatorias continuas utilizadas en teoría de probabilidad y estadística.
4	F	Las áreas tabuladas de la distribución normal de probabilidad son herramientas prácticas que simplifican los cálculos relacionados con la distribución normal, lo que facilita tanto el análisis de datos simples como el manejo de situaciones más complejas en diversos contextos estadísticos.
5	V	La variable aleatoria normal estándar es una variable aleatoria continua que sigue una distribución normal con una media de cero y una desviación estándar de uno.
6 Ejercicio 1		La probabilidad de que una muestra determinada de 1 ml contenga más de 100 bacterias es aproximadamente 0,0475, o el 4,75 %.
7 Ejercicio 2		La probabilidad de que la temperatura en un mes dado esté por encima de 30 °C es del 4,75 %.
8 Ejercicio 3 a.		La probabilidad de encontrar más de 30 especies de abejas en un sitio de muestreo es aproximadamente 84,13 %.
9 Ejercicio 3 b.		La probabilidad de encontrar entre 20 y 30 especies de abejas en un sitio de muestreo es aproximadamente $0,8413 - 0,1587 = 0,6826$ o el 68.26 %.
10 Ejercicio 3 c.		Aproximadamente el 2,28 % de los sitios de muestreo tendrán 15 o menos especies de abejas.

[Ir a la autoevaluación](#)

Autoevaluación 11

Pregunta	Respuesta	Retroalimentación
1	V	En estadística, la población se refiere al grupo completo que se quiere estudiar, mientras que la muestra se refiere a un subconjunto representativo de esa población. Esto es fundamental para poder realizar inferencias sobre la población con base en la información recopilada de la muestra.
2	F	La estadística y las distribuciones muestrales son herramientas fundamentales para obtener información sobre la variabilidad y distribución de características en un ecosistema forestal restaurado, lo que ayuda en la toma de decisiones y la evaluación del éxito del proyecto.
3	F	El teorema del límite central es relevante no solo para muestras pequeñas, sino también para muestras grandes. Este teorema establece que, independientemente de la forma de la distribución poblacional, la distribución de las medias de las muestras se aproxima a una distribución normal a medida que el tamaño de la muestra aumenta. Esto es fundamental en inferencia estadística, ya que permite hacer suposiciones sobre la distribución de las medias muestrales y realizar pruebas de hipótesis y construcción de intervalos de confianza.
4	Ejercicio 1	Aunque el muestreo de cada décima persona puede parecer aleatorio, este método puede generar sesgos significativos en el estudio. Por ejemplo, es probable que este enfoque subrepresente a ciertos grupos demográficos que tienden a visitar parques urbanos en diferentes momentos o días. Además, puede ignorar las preferencias de aquellos que no visitan los parques con la misma frecuencia. Por lo tanto, la muestra obtenida puede no ser representativa de la población en general y puede sesgar los resultados del estudio sobre las preferencias de actividades recreativas en áreas verdes. Para obtener una muestra más representativa, sería preferible utilizar un método de muestreo aleatorio simple que garantice que cada individuo de la población tenga la misma probabilidad de ser seleccionado para participar en el estudio.



Pregunta	Respuesta	Retroalimentación
5 Ejercicio 2 a.		<p>El universo o población está compuesto por todos los niños que asisten a la escuela ubicada en el centro de la ciudad y que cumplen con los criterios de elegibilidad del estudio (residir en la ciudad, asistir regularmente a la escuela y tener consentimiento de sus padres). Se selecciona una muestra de 50 niños de entre 8 y 12 años de edad que cumplan con los criterios de elegibilidad del estudio. El estudio se lleva a cabo durante nueve meses para capturar posibles variaciones estacionales en la calidad del aire y en la salud respiratoria de los niños. Durante el periodo de estudio, los investigadores monitorean la calidad del aire en áreas cercanas a la escuela utilizando equipos especializados. Además, recolectan datos sobre la frecuencia de síntomas respiratorios en los niños, como tos, dificultad para respirar o sibilancias, a través de cuestionarios, entrevistas o registros médicos.</p>
6 Ejercicio 2 b.		<p>Para ello, se puede utilizar un muestreo aleatorio simple o algún otro método de muestreo probabilístico, asegurando que todos los niños elegibles tengan la misma probabilidad de ser seleccionados.</p>
7 Ejercicio 2 c.		<p>Sí, se pueden hacer inferencias válidas utilizando los resultados de este estudio. Si el estudio está bien diseñado y ejecutado, y si se tienen en cuenta las limitaciones y consideraciones adicionales, es posible hacer inferencias válidas sobre la relación entre la exposición a la contaminación del aire y los efectos en la salud de los niños en el área urbana en cuestión.</p>
8 Ejercicio 2 d.		<p>En lugar de seleccionar aleatoriamente 50 niños de entre 8 y 12 años de la población general de la escuela, podría utilizar un enfoque de muestreo estratificado. Esto implica dividir la población de la escuela en grupos o estratos basados en características importantes, como la edad, el género, el nivel socioeconómico o la distancia desde las fuentes de contaminación del aire. Luego, se seleccionaría una muestra aleatoria de niños de cada estrato para garantizar una representación adecuada de todas las subpoblaciones relevantes. También, para evaluar más efectivamente los efectos de la contaminación del aire en la salud de los niños, podría incluir un grupo de control de niños que asisten a una escuela en un área con niveles más bajos de contaminación del aire. Comparar los resultados entre el grupo expuesto y el grupo de control ayudaría a identificar los efectos específicos de la contaminación del aire.</p>



9 Ejercicio
3
a.

- Cambios en la temperatura ambiente o la humedad relativa pueden afectar la lectura de los termómetros digitales y, por lo tanto, las mediciones de la temperatura del agua.
- Pequeñas desviaciones en la calibración de los termómetros digitales pueden conducir a diferencias en las lecturas de temperatura entre dispositivos, incluso de la misma marca y modelo.
- La temperatura del agua puede variar con la profundidad, especialmente en ríos donde el agua más profunda puede tener una temperatura ligeramente diferente a la de la superficie.
- La velocidad del flujo de agua puede influir en la distribución y mezcla de la temperatura dentro del río, lo que puede afectar las mediciones de temperatura en diferentes puntos a lo largo del curso del agua.
- La temperatura del agua puede variar a lo largo del día debido a cambios en la radiación solar, la actividad biológica y otros factores temporales.
- Descargas de aguas residuales industriales o de plantas de energía pueden introducir agua caliente en el río, afectando localmente la temperatura del agua en ciertos puntos de muestreo.
- La presencia de sombra de la vegetación circundante puede afectar la temperatura del agua en áreas específicas del río, creando gradientes térmicos que podrían influir en las mediciones.

10 Ejercicio
3
b.

- Para reducir la cantidad de variabilidad en el promedio de mediciones, se debe utilizar un número grande de repeticiones. Esto se debe a que, con un mayor número de repeticiones, se obtiene una estimación más precisa del verdadero valor de la temperatura del agua en cada punto del río. Al tomar el promedio de un mayor número de mediciones, se reduce el impacto de cualquier error de medición individual y se mejora la precisión del resultado.
- Cuando se promedian un gran número de mediciones, cualquier variabilidad aleatoria en las mediciones individuales tiende a “promediarse” y su efecto se reduce. Por lo tanto, cuanto más grande sea el número de repeticiones, más confiable será el promedio muestral y más representativo será de la verdadera temperatura del agua en ese punto del río.

[Ir a la autoevaluación](#)



Autoevaluación 12

Pregunta	Respuesta	Retroalimentación
1	V	Los reportes de investigación son una forma efectiva de comunicar hallazgos, conclusiones y recomendaciones derivadas de la investigación.
2	F	Los informes de investigación son cruciales para el avance del conocimiento en el campo ambiental al compartir información sobre los resultados y las implicaciones de la investigación.
3	F	Los informes de investigación proporcionan información valiosa que puede ayudar en la toma de decisiones para la conservación y gestión sostenible del ambiente al compartir hallazgos y recomendaciones basadas en evidencia científica.
4	F	Los reportes de investigación son herramientas clave para la difusión de información científica entre la comunidad académica y los responsables de políticas ambientales, lo que promueve una mejor comprensión de los problemas ambientales y posibles soluciones.
5	F	Los reportes de investigación pueden promover la conciencia y la acción en beneficio del ambiente al difundir información científica precisa y relevante que puede inspirar cambios de comportamiento y políticas en la sociedad.
6	F	El reporte de investigación es un documento que necesita ser cuidadosamente planificado y estructurado para cumplir con su propósito de comunicar efectivamente los resultados de la investigación.
7	V	La definición clara de los usuarios determinará el formato, la naturaleza y la extensión del informe, lo que garantiza que sea relevante y útil para su audiencia prevista.
8	F	Los reportes de investigación pueden presentarse tanto en contextos académicos como no académicos, y el contexto influye en su diseño y contenido.
9	F	La estructura común de un reporte académico incluye una serie de secciones adicionales como la portada, índices, resumen, revisión de la literatura, referencias y apéndices, además de las secciones mencionadas.
10	V	La revisión de la literatura incluye una sección que aborde la distribución de investigaciones relacionadas con los Objetivos de Desarrollo Sostenible (ODS), según el esquema proporcionado.

Pregunta Respuesta Retroalimentación

Ir a la autoevaluación





5. Referencias bibliográficas

Douglas, M. C. (2003). *Diseño y análisis de experimentos*. Limusa Wiley.

Domínguez, J. D., & Tostado, E. C. (2016). *Diseño de experimentos: Estrategias y análisis en ciencias e ingenierías*. Alpha Editorial.

Hernández Sampieri, R., Fernández Collado, C., & Baptista Lucio, P. (2018). *Metodología de la investigación* (Vol. 4, pp. 310-386). México: McGraw-Hill Interamericana.

Mendenhall W, Beaver R & Beaver B. (2015). *Introducción a la probabilidad y estadística*. 14a edición. México, D.F.: Cengage Learning.

Bozorg-Haddad, O., Delpasand, M., & Loaiciga, H. A. (2021). Water quality, hygiene, and health. In *Economical, Political, and Social Issues in Water Resources* (pp. 217-257). Elsevier.

Croijmans, L., De Jong, J. F., & Prins, H. H. T. (2021). Oxygen is a better predictor of macroinvertebrate richness than temperature—a systematic review. *Environmental Research Letters*, 16(2), 023002.

Gavilánez-Luna, F. (2021). *Diseños y análisis estadísticos para experimentos agrícolas*. Ediciones Díaz de Santos.

Gutiérrez-Pulido, H., & Vara-Salazar, R. (2012). *Análisis y diseño de experimentos* (3a ed.). McGraw-Hill.

Melo, L. L. O., Melo, S., & López, L. (2007). *Diseño de experimentos: Métodos y aplicaciones*. Editorial Universidad Nacional de Colombia.

Sun, L., Song, F., Liu, S., Cao, Q., Liu, F., & Zhu, X. (2018). Integrated agricultural management practice improves soil quality in Northeast China. *Archives of Agronomy and Soil Science*, 64(14), 1932-1943.

La profe mónica. (18 de febrero de 2023). *¡Superfórmula! Para redactar objetivos / objetivos generales / objetivos específicos* [Video]. YouTube. Obtenido de <https://www.youtube.com/watch?v=DQlp0DvPBwU>

Teocom. (18 de junio de 2023). *Ejemplos de alcances de una investigación* [Video]. YouTube. Obtenido de <https://www.youtube.com/watch?v=zZ3hzUVIN6M>

Reskala, F. (5 de julio de 2021). *Diseño experimental y no experimental explicados* [Video]. YouTube. https://www.youtube.com/_watch?v=1065eCE69ws

Martínez, C. (22 de enero de 2021). *Diagrama de caja y bigotes - Estadística en Excel* [Video]. YouTube. https://www.youtube.com/_watch?v=fpRbZr-t8bk

Gómez, A. (15 de septiembre de 2018). *Cuartiles introducción / Qué son y como encontrarlos en datos sin agrupar* [Video]. YouTube. <https://www.youtube.com/watch?v=suSz9RXFNTs>

Trabaja con datos. (27 de junio de 2023). *¡Tutorial gráfico de dispersión y correlación en excel facil! ¡Plantilla gratis!* [Video]. YouTube. Obtenido de <https://www.youtube.com/watch?v=Df8BkwIzZ3Q>

Matemóvil. (6 de febrero de 2019). *Experimento aleatorio, espacio muestral, evento o suceso y probabilidades* [Video]. YouTube. Obtenido de <https://www.youtube.com/watch?v=fTIS83G7aC8>

Matemóvil. (3 de febrero de 2020). *Función de probabilidad de variable aleatoria discreta / Ejercicio 1* [Video]. YouTube. Obtenido de <https://www.youtube.com/watch?v=9NicRz7xPU0>