

ESTADÍSTICA DESCRIPTIVA

Aplicada en la Arquitectura e Ingeniería Civil



Mgtr. Eliana Contreras Jordán



Universidad Laica
VICENTE ROCAFUERTE
de Guayaquil

Estadística Descriptiva aplicada en la Arquitectura e Ingeniería Civil

Fecha de publicación: 10 de octubre de 2025

ISBN: 978-9942-617-16-3

Derechos de autor: GYE-015420

Hecho el Depósito Legal

Autora

Eliana Noemí Contreras Jordán. [ORCID](#)

Filiación institucional

La autora ejercía la calidad de docente de la Universidad Laica VICENTE ROCAFUERTE de Guayaquil al momento de la escritura del libro.

De esta edición

Universidad Laica VICENTE ROCAFUERTE de Guayaquil (ULVR), 2025

Av. de Las Américas #70, frente al Cuartel Modelo, Guayaquil, Ecuador

PBX: (00-593-4) 259 6500. www.ulvr.edu.ec

Revisión de pares

El libro *Estadística Descriptiva aplicada en la Arquitectura e Ingeniería Civil*, fue arbitrado por el Departamento de Investigación Científica, Tecnología e Innovación de la ULVR bajo la metodología doble par ciego (*double blind peer review*). Los revisores son externos a la ULVR.

Línea de investigación: Territorio, medioambiente y materiales innovadores para la construcción.

Campo Detallado: 2-37A Construcción e ingeniería civil.

Para referenciar el libro, utilice el siguiente formato, de acuerdo a las Normas APA 7ª edición:

Contreras, E. (2025, 10 de octubre). *Estadística Descriptiva aplicada en la Arquitectura e Ingeniería Civil*. Editorial ULVR.

Palabras clave: [Estadística](#), [Arquitectura](#), [Ingeniería Civil](#), [Diseño estructural](#), [Diseño de proyecto](#).

Keywords: [Statistics](#), [Architecture](#), [Civil engineering](#), [Structural design](#), [Project design](#).

Consejo Editorial de la Universidad Laica VICENTE ROCAFUERTE de Guayaquil

- Ph.D Aimara Rodríguez Fernández, Rectora
- Ph.D Óscar Parada Gutiérrez, Vicerrector Académico de Investigación, Grado y Posgrado
- Mgtr. Alex Salvatierra Espinoza, Vicerrector Administrativo
- Mgtr. Jessica Aroca Clavijo, Decana de la Facultad de Administración
- Mgtr. Carlos Pérez Leyva, Decano de la Facultad de Ciencias Sociales y Derecho
- Ph.D Luis Manzano Díaz, Decano de la Facultad de Educación
- Ph.D Marcial Calero Amores, Decano de la Facultad de Ingeniería, Industria y Construcción
- Mgtr. Alfredo Aguilar Hinojosa, Director del Departamento de Marketing y Relaciones Internacionales
- Mgtr. Patricia Navarrete Zavala, Editorialista ULVR



Editado por: Editorial ULVR

edilaica@ulvr.edu.ec

PBX: (00-593-4) 259 6500, extensión 195



ESTADÍSTICA DESCRIPTIVA

Aplicada en la Arquitectura e Ingeniería Civil

Mgtr. Eliana Contreras Jordán



Universidad Laica
VICENTE ROCAFUERTE
de Guayaquil

ÍNDICE

Introducción

Capítulo 1

Introducción a la Estadística en Ingeniería Civil

1.1.	Evaluación de Riesgos y Diseño Estructural	12
1.2.	Análisis de Carga y Estructuras	12
1.3.	Control de Calidad y Aseguramiento de la Calidad.....	13
1.4.	Planificación y Programación de Proyectos	13
1.5.	Sostenibilidad y Eficiencia Energética	13
1.6.	Definición de Estadística	14
1.6.1.	La estadística descriptiva	15
1.6.2.	La estadística Inferencial.....	15
1.7.	Conceptos fundamentales.....	16
1.7.1.	Población	16
1.7.2.	Censo	17
1.7.3.	Muestra	18
1.7.4.	Variable.....	19
1.7.5.	Dato/Datos	20
1.7.6.	Parámetro	21
1.7.7.	Estadístico	22
1.7.8.	Hipótesis	24
1.7.9.	Prueba de Hipótesis.....	25
1.7.10.	Muestreo Aleatorio	27
1.7.11.	Error Estándar	28
1.7.12.	Test de Significancia	29
1.7.13.	Nivel de Significancia.....	30
1.7.14.	Error Tipo I	31
1.7.15.	Error Tipo II	32
1.7.16.	Diseño Experimental	33
1.8.	Tipos de datos	35
1.8.1.	Datos cualitativos	35
1.8.2.	Clasificación datos cualitativos	35
1.8.2.1.	Datos Cualitativos Nominale.....	35
1.8.2.2.	Datos Cualitativos Ordinales	36
1.8.3.	Datos cuantitativos	38

1.8.4.	Clasificación datos cuantitativos	39
1.8.4.1.	Datos Cuantitativos Discretos:	39
1.8.4.2.	Datos Cuantitativos Continuos.....	39
1.9.	Tipos de medición	40
1.9.1.	Nivel de Medición Nominal	40
1.9.2.	Nivel de Medición Ordinal.....	41
1.9.3.	Nivel de Medición de Intervalo	42
1.9.4.	Nivel de Medición de Razón	43
1.10.	El proceso de análisis estadístico	45
1.10.1.	Recopilación de Datos	45
1.10.2.	Organización de Datos.....	45
1.10.3.	Limpieza de Datos.....	45
1.10.4.	Exploración de Datos	46
1.10.5.	Análisis Estadístico Descriptivo.....	46
1.10.6.	Análisis Estadístico Inferencial	46
1.10.7.	Interpretación de resultados.....	46
1.10.8.	Presentación de Resultados	47
1.10.9.	Toma de Decisiones	47
1.11.	Ejemplos de aplicación del proceso de análisis estadístico	47
1.11.1.	Guía de trabajo 1: Evaluación de Riesgos Sísmicos en un Puente de Importancia Crítica.....	48
1.11.2.	Guía de trabajo 2: Evaluación de Riesgos en una Represa Hidroeléctrica	51
	Ejercicios a desarrollar del Capítulo 1	54
	Conclusiones del capítulo 1	57

Capítulo 2

Recopilación y Organización de Datos

2.1.	Concepto de recolección de datos.....	59
2.2.	Técnica de recolección de datos	60
2.2.1.	Encuestas.....	60
2.2.2.	Experimentos.....	60
2.2.3.	Observación Directa	60
2.2.3.1.	Preparación para la Observación.....	61
2.2.4.	Registros y Bases de Datos.....	61
2.2.5.	Entrevistas	61
2.2.6.	Cuestionarios en Línea	61
2.2.7.	Análisis de Contenido.....	61
2.2.8.	Sensores y Dispositivos Inteligentes	61
2.2.9.	Grupos de Discusión	62
2.3.	Instrumentos de Recolección de Datos en Estadística	62

2.3.1.	Cuestionarios y Encuestas	62
2.3.2.	Entrevistas	62
2.3.3.	Observación Directa	63
2.3.4.	Registro de Datos	63
2.3.5.	Escalas de Medición	63
2.3.6.	Sensores y Dispositivos Inteligentes	63
2.3.7.	Grupos de Discusión	63
2.3.8.	Análisis de Contenido.....	64
2.4.	Organización de los datos.....	64
2.4.1.	Tablas de Frecuencias	64
2.4.2.	Gráficos y Diagramas	65
2.4.3.	Matrices de Datos.....	66
2.4.4.	Clasificación de Tablas de Frecuencias.....	67
2.4.4.1.	Tabla de Frecuencias para Datos No Agrupados o Discretos simples	67
2.4.4.2.	Tabla de Frecuencias para Datos Agrupados en Intervalos o Continuos.....	68
2.4.4.3.	Tabla de Frecuencias para Datos Categóricos.....	68
2.4.4.4.	Tabla de Frecuencias Cruzadas o de Contingencia	69
2.5.	Componentes de una Tabla de Frecuencias.....	71
2.5.1.	Clases o Intervalos.....	71
2.5.2.	Frecuencia Absoluta (fa)	71
2.5.3.	Frecuencia Relativa (fr).	71
2.5.4.	Frecuencia Acumulada (Fa)	72
2.5.5.	Frecuencia Relativa Acumulada (Fr)	72
2.6.	Creación de tablas de frecuencias para datos simples.	72
2.7.	Ejercicio de construcción de tablas de frecuencias para datos simples	73
2.8.	Creación de tablas de frecuencias para datos agrupados	75
2.9.	Ejercicio de construcción de tablas de frecuencias para datos agrupados.....	78
	Ejercicios a desarrollar del Capítulo 2	80
	Conclusiones del capítulo 2	82

Capítulo 3

Medidas de Tendencia Central

3.1.	Introducción.....	85
3.2.	Media Aritmética.....	85
3.2.1.	Pasos para Calcular la Media de Datos Simples.....	86
3.2.2.	Pasos para Calcular la Media de Datos Agrupados	86
3.2.3.	Ejemplos para calcular la Media Aritmética datos simples	87

3.2.4.	Ejemplos para calcular la Media Aritmética datos agrupados.....	89
3.3.	Mediana	92
3.3.1.	Pasos para Calcular la Mediana de Datos Simples	92
3.3.2.	Pasos para Calcular la Mediana de Datos Agrupados	93
3.3.3.	Ejemplo de Mediana.....	93
3.4.	Moda.....	96
3.4.1.	Ejemplo de Moda	96
3.5.	Usos y aplicaciones de las medidas de tendencia central	99
3.5.1.	Uso de la Media en Ingeniería Civil y Arquitectura.....	99
3.5.2.	Uso de la Mediana en Ingeniería Civil y Arquitectura.....	100
3.5.3.	Uso de la Moda en Ingeniería Civil y Arquitectura	101
	Ejercicios a desarrollar del Capítulo 3	103
	Conclusiones del capítulo 3	110

Capítulo 4

Medidas de Dispersión

4.1.	Introducción.....	113
4.2.	Rango.....	113
4.3.	Varianza	114
4.4.	Desviación estándar	117
4.5.	Coeficiente de variación	119
4.6.	Caso de Estudio Simulado: Análisis de Costos en Proyecto de Desarrollo de Viviendas	123
	Ejercicios a desarrollar del Capítulo 4.....	126
	Conclusiones del capítulo 4	129

Capítulo 5

Medidas de Tendencia Central

5.1.	Introducción	132
5.2.	Diagramas de barras simples	133
5.2.1	Características	133
5.2.2	Ventajas	133
5.3.	Diagramas de Barras compuestos	134
5.3.1	Características	135
5.3.2	Ventajas	135
5.4.	Hitogramas.....	136
5.4.1	Características	137
5.4.2	Ventajas	137
5.5.	Gráficos Circular, Sector o Pastel.....	140
5.5.1	Características	140
5.5.2	Ventajas	140

5.6.	Polígono de frecuencias acumuladas	142
5.6.1	Características	143
5.6.2	Ventajas	143
Conclusiones del capítulo 5.....		146
Conclusiones		147
Referencias		149
Respuestas a los ejercicios		156

INTRODUCCIÓN

En el ámbito de la Arquitectura e Ingeniería civil, donde cada puente, carretera, edificio y estructura es un testimonio tangible del ingenio humano, la toma de decisiones precisas y fundamentadas en datos es necesaria. La Ingeniería Civil es un campo multidisciplinario que abarca desde la planificación y el diseño hasta la construcción y la gestión de infraestructuras que dan forma a nuestras comunidades y facilitan la vida moderna. En este contexto, las estadísticas se revelan como una herramienta poderosa que impulsa la toma de decisiones informadas, la mejora de la eficiencia y la gestión eficaz de recursos limitados.

La estadística descriptiva es una herramienta fundamental en el análisis de datos, proporcionando métodos y técnicas para resumir y describir la información recopilada en estudios y proyectos. En los campos de la arquitectura y la ingeniería civil, la aplicación de estas técnicas es fundamental para la toma de decisiones informadas, la planificación eficiente y la ejecución precisa de proyectos complejos.

En la arquitectura, la estadística descriptiva permite a los profesionales analizar datos sobre materiales, costos, tiempos de construcción y preferencias de los usuarios, entre otros. Por ejemplo, al evaluar la distribución de los costos en diferentes proyectos, los arquitectos pueden identificar tendencias y tomar decisiones más económicas y eficientes en futuros diseños. Asimismo, el análisis de encuestas y estudios de satisfacción del cliente ayuda a ajustar los diseños para satisfacer mejor las necesidades de los usuarios finales.

En la ingeniería civil, la estadística descriptiva es igualmente vital. Los ingenieros utilizan técnicas estadísticas para analizar datos sobre la

resistencia de materiales, la durabilidad de estructuras y la eficiencia de los procesos constructivos. Los histogramas y gráficos de barras permiten visualizar la distribución de los datos de pruebas de materiales, mientras que los polígonos de frecuencias acumuladas (ojivas) facilitan la comprensión de la acumulación de riesgos o costos a lo largo del tiempo. Este libro tiene como objetivo proporcionar una guía exhaustiva y práctica sobre el uso de la estadística descriptiva en estos campos. A través de ejemplos específicos, estudios de caso y ejercicios aplicados, se pretende equipar a los lectores con las habilidades necesarias para utilizar herramientas estadísticas en la resolución de problemas reales en la práctica profesional.

El contenido está estructurado para ofrecer tanto una comprensión teórica de los conceptos estadísticos como su aplicación práctica. Desde la recolección y organización de datos hasta su análisis e interpretación, este libro aborda todas las etapas del proceso estadístico. Se incluyen discusiones detalladas sobre gráficos de barras, histogramas, gráficos de sectores y polígonos de frecuencias acumuladas, destacando su relevancia y utilidad en contextos arquitectónicos y de ingeniería civil.

Capítulo 1

INTRODUCCIÓN A LA ESTADÍSTICA EN INGENIERÍA CIVIL

En este capítulo, se presenta la importancia de la estadística en la ingeniería civil y su aplicación en el diseño, construcción y gestión de infraestructuras. Se analiza cómo el uso de datos contribuye a la toma de decisiones, la eficiencia de los proyectos y la seguridad estructural. También se abordan métodos estadísticos que facilitan la planificación y evaluación en este campo.



La estadística desempeña un papel fundamental, ya que esta disciplina se enfrenta constantemente a la complejidad de diseñar, construir y gestionar infraestructuras que son esenciales para nuestras comunidades. En este capítulo, exploraremos la relevancia de las estadísticas y cómo su aplicación adecuada puede marcar la diferencia en la toma de decisiones, la eficiencia de los proyectos y la seguridad de las estructuras.

1.1. Evaluación de Riesgos y Diseño Estructural

La evaluación de riesgos es una parte crítica de la ingeniería civil, y las estadísticas son la base de este proceso. Los ingenieros civiles utilizan métodos estadísticos para modelar eventos extremos, como terremotos, inundaciones y vientos fuertes, y calcular la probabilidad de que ocurran durante la vida útil de una estructura. Esto es esencial para diseñar infraestructuras resistentes que puedan soportar las condiciones más adversas.

1.2. Análisis de Carga y Estructuras

En la ingeniería civil, es fundamental comprender cómo las cargas afectan a las estructuras. Las estadísticas permiten a los ingenieros analizar y prever las cargas máximas que las infraestructuras pueden soportar, lo que es clave para el diseño de puentes, edificios y otras estructuras. Este análisis estadístico, al involucrar la recopilación y evaluación de datos históricos y actuales, ayuda a identificar los comportamientos esperados bajo diferentes condiciones de carga.

Además, las estadísticas son esenciales en la evaluación de riesgos, como en la predicción de fallos estructurales, a través de técnicas que modelan las propiedades de materiales y las cargas dinámicas. Un ejemplo de ello es el trabajo realizado por Nguyen et al. (2020), quienes utilizaron el método de suavizado interrumpido (GSM) combinado con redes neuronales convolucionales (CNN) para detectar daños en puentes.

Este enfoque integrado, que combina estadísticas y análisis de datos dinámicos, no solo permite prever el comportamiento de las infraestructuras ante cargas previstas, sino también identificar de

manera temprana los daños estructurales, mejorando la seguridad y eficiencia del mantenimiento.

1.3. Control de Calidad y Aseguramiento de la Calidad

La seguridad y la calidad son de suma importancia en la ingeniería civil. Las estadísticas desempeñan un papel esencial en el control de calidad de los materiales utilizados en la construcción, así como en la inspección y supervisión de proyectos. El control estadístico de procesos (CEP) es una herramienta clave para garantizar que los estándares de calidad se cumplan de manera consistente. Según Cobb (2024), los modelos de CEP modernos no solo permiten monitorear la calidad de manera eficiente, sino que también optimizan los costos al ajustar los intervalos y tamaños de muestreo según las condiciones del proceso. Además, al integrar variaciones no constantes en los procesos, el CEP contribuye significativamente a minimizar las desviaciones y a mantener la calidad del producto final.

1.4. Planificación y Programación de Proyectos

La planificación eficiente es esencial en la ingeniería civil, donde los proyectos pueden ser enormes y costosos. Las estadísticas permiten a los ingenieros realizar análisis de tiempo y costos, lo que ayuda a programar proyectos de manera óptima y asignar recursos de manera eficiente. Además, la gestión de riesgos se beneficia de técnicas estadísticas para evaluar y mitigar retrasos y desviaciones presupuestarias. Por ejemplo, Bai (2023) propuso el uso de algoritmos de árboles de decisión para mejorar la precisión en la predicción de costos durante el ciclo de vida de proyectos de ingeniería civil. Este enfoque permite un análisis más eficiente y preciso, optimizando los recursos y reduciendo los riesgos asociados con desviaciones presupuestarias.

1.5. Sostenibilidad y Eficiencia Energética

En una época en la que la sostenibilidad es una preocupación global, las estadísticas son esenciales para evaluar el impacto ambiental de

las infraestructuras y diseñar soluciones eficientes desde el punto de vista energético. El análisis estadístico de datos relacionados con el consumo de energía, las emisiones de carbono y el uso de recursos naturales ayuda a los ingenieros a tomar decisiones que reduzcan el impacto ambiental de los proyectos de construcción. Por ejemplo, Sun (2023) analizó cómo la infraestructura digital mejora la eficiencia energética y ambiental mediante modelos estadísticos avanzados, destacando que la tecnología digital y la investigación en innovación energética son fundamentales para aumentar la sostenibilidad en los proyectos de infraestructura.

1.6. Definición de Estadística

La estadística es una disciplina matemática y científica fundamental para la recopilación, organización, análisis, interpretación y presentación de datos. Según Zorić (2021), su objetivo principal es extraer información útil para la toma de decisiones informadas, lo que la convierte en una herramienta esencial en ámbitos como la investigación científica, la industria, la medicina y la gestión empresarial. La estadística no solo se limita al análisis de datos, sino que también aborda la inferencia estadística, lo que permite a los investigadores hacer generalizaciones a partir de muestras y probar hipótesis con rigor matemático.

Además, la estadística tiene aplicaciones diversas y críticas en el mundo actual. Por ejemplo, Williams y Hooten (2016) destacan cómo se utiliza en la gestión de decisiones complejas bajo incertidumbre, integrando la teoría estadística con herramientas prácticas como la estimación puntual y los intervalos de confianza. Asimismo, Harrison y Tamaschke (1993) señalan su rol central en la toma de decisiones en negocios y economía, donde los análisis estadísticos permiten a las empresas mejorar la calidad, la productividad y las estrategias de mercado mediante la comprensión de variaciones y patrones en los datos.

En el ámbito de la ciencia, Gouveia-Oliveira (1996) resalta que la estadística es esencial para la medicina moderna, desde la identificación de factores de riesgo hasta la elección de tratamientos

óptimos. Este enfoque interdisciplinario demuestra que, lejos de ser una disciplina aislada, la estadística se posiciona como un lenguaje universal para describir fenómenos y guiar acciones en contextos variados, desde proyectos ecológicos hasta decisiones empresariales. La estadística se divide en estadística descriptiva y estadística inferencial.

1.6.1. La estadística descriptiva

Se centra en resumir y presentar datos de manera informativa, utilizando medidas de tendencia central, medidas de dispersión y representaciones gráficas. Es la encargada de recolectar, organizar, resumir y presentar datos de manera informativa. Su objetivo es describir las características básicas de un conjunto de datos, utilizando medidas como la media, la mediana, la moda, la varianza y los gráficos.

La estadística descriptiva ayuda a los ingenieros y arquitectos a entender y resumir grandes cantidades de datos. Por ejemplo, si se requiere evaluar la calidad del concreto en varios edificios, usaría la estadística descriptiva para calcular un promedio de la resistencia del concreto y así saber si la mayoría de los edificios cumplen con los estándares de seguridad.

Ejemplo:

Un ingeniero civil está analizando la calidad del concreto utilizado en la construcción de un conjunto de edificios. Se recolectan muestras de concreto de 50 edificios para medir su resistencia a la compresión. La estadística descriptiva se utiliza para calcular la resistencia media de estas muestras, determinar la desviación estándar (que indica la variabilidad en la resistencia), y presentar los resultados en gráficos para observar la distribución de las resistencias en todas las muestras.

1.6.2. La estadística Inferencial

Son métodos estadísticos utilizados para hacer inferencias o generalizaciones sobre una población a partir de una muestra. Es una

rama de la estadística que se encarga de hacer generalizaciones, predicciones o inferencias sobre una población basándose en una muestra de datos. Esto incluye estimar parámetros, realizar pruebas de hipótesis y crear modelos predictivos.

En lugar de analizar todos los datos posibles, los ingenieros civiles pueden usar una pequeña parte de los datos disponibles para hacer predicciones o tomar decisiones informadas sobre proyectos más grandes. Por ejemplo, un ingeniero podría evaluar solo un conjunto limitado de puentes para inferir la durabilidad de un nuevo material en todos los puentes de una región.

Ejemplo:

Un ingeniero civil está evaluando la resistencia al desgaste de un nuevo tipo de pavimento asfáltico. En lugar de probar cada tramo de carretera, el ingeniero toma una muestra representativa de 100 secciones de carretera de un total de 1,000 secciones construidas con este pavimento. Usando la estadística inferencial, el ingeniero puede estimar la resistencia al desgaste promedio de todo el pavimento basado en esta muestra, y determinar con un nivel de confianza si es necesario mejorar el material antes de su aplicación masiva.

1.7. Conceptos fundamentales

1.7.1. Población

Es la colección completa de todos los elementos (puntuaciones, personas, mediciones, etcétera) a estudiar. Se dice que la colección es completa, pues incluye a todos los sujetos que se estudiarán. Este conjunto puede incluir personas, objetos, eventos, o cualquier unidad de análisis que se esté investigando. Según Sahu (2016), el concepto de población es fundamental en estadística, ya que permite caracterizar grupos definidos por características comunes, ya sea en contextos finitos, como estudiantes en una universidad, o infinitos, como estrellas en una galaxia. Esta distinción es esencial para estructurar estudios estadísticos y diseñar muestreos adecuados.

Ejemplos:

- 1. En un proyecto para mejorar la eficiencia energética, un arquitecto puede estudiar todos los edificios de oficinas construidos entre 2010 y 2020 en una ciudad específica. Aquí, la población incluiría todos esos edificios, y el arquitecto podría seleccionar una muestra representativa para evaluar el rendimiento energético y luego aplicar las conclusiones a la totalidad de la población.*
- 2. Se requiere analizar la durabilidad del concreto utilizado en todos los puentes de una ciudad. En este caso, el ingeniero civil considerará la población como el total de todos los puentes construidos con ese tipo de concreto en la ciudad. Si hay 200 puentes, entonces la población sería esos 200 puentes. El ingeniero podría tomar una muestra de 20 puentes (una muestra es un subconjunto de la población) para realizar pruebas de durabilidad y usar esos resultados para hacer inferencias sobre la durabilidad del concreto en todos los puentes de la población.*

1.7.2. Censo

Es la colección de datos de cada uno de los miembros de la población. Es un método de recolección de datos donde se obtiene información de todos los elementos de una población completa, en lugar de solo una muestra.

Un censo es un enfoque exhaustivo para obtener información sobre cada elemento de interés, como todos los edificios en una ciudad o todas las carreteras en una región, para tomar decisiones más informadas y completas.

Según Sahai (1988), el censo es un proceso integral que incluye la recopilación y publicación de datos demográficos, socioeconómicos y de otros tipos relevantes de todos los individuos de una región, siendo fundamental para entender estructuras poblacionales y facilitar la toma de decisiones basadas en datos.

Ejemplos:

- 1. Se evalúa el estado de las carreteras en una región específica después de un terremoto. En lugar de revisar solo algunas carreteras, realiza un censo, inspeccionando cada una de las carreteras en esa región*

para identificar daños. Esto permite obtener un panorama completo de la situación, y determinar con precisión qué zonas necesitan reparación inmediata. En ingeniería civil, un censo es como hacer un inventario completo de algo que se quiere estudiar, como todas las carreteras de una ciudad después de un evento como un terremoto, asegurándose de no dejar ninguna fuera del análisis.

- 2. Para diseñar un plan de renovación en un conjunto de edificios históricos en una ciudad, se realiza un censo para catalogar todos los edificios que serán parte del proyecto, incluyendo detalles sobre su estado actual, los materiales usados y cualquier daño estructural. Esto permite al arquitecto entender el estado de todos los edificios y planificar las renovaciones de manera integral. En arquitectura, un censo podría implicar revisar y documentar todos los edificios en un área específica antes de iniciar un proyecto de renovación, para asegurarse de que se tienen en cuenta todos los aspectos de cada edificio.*

1.7.3. Muestra

Es un subconjunto de miembros seleccionados de una población. Este subconjunto se elige para representar a toda la población, permitiendo realizar inferencias y análisis sin necesidad de estudiar cada elemento de la población completa. Una muestra es como tomar una pequeña parte de un todo para analizarla y entender cómo es el conjunto completo, ya sea una red de carreteras, un complejo de oficinas o una serie de edificios históricos. Es una forma eficiente y práctica de obtener información relevante sin necesidad de revisar cada elemento individualmente. Según Saldanha y Thompson (2002), una muestra es una herramienta que, mediante el proceso de muestreo, permite razonar sobre la variabilidad y distribución de las características de la población total, haciendo posible extrapolar conclusiones fiables.

Ejemplos:

- 1. Supongamos que un ingeniero civil quiere evaluar la durabilidad del pavimento en una ciudad. En lugar de analizar cada carretera, selecciona una muestra de tramos de carretera que incluye*

diferentes zonas (urbana, rural), tipos de pavimento y niveles de tráfico. Esta muestra le permitirá inferir el estado general del pavimento en toda la ciudad. En ingeniería civil, una muestra permite a los ingenieros hacer un análisis representativo sin tener que revisar cada kilómetro de carretera. Seleccionan tramos específicos que reflejan las diferentes condiciones y características de toda la red vial.

- 2. Un arquitecto está diseñando un estudio sobre la iluminación natural en un complejo de oficinas. En lugar de medir la luz en cada oficina, selecciona una muestra de oficinas que incluye diferentes orientaciones, tamaños de ventanas y alturas de piso. Esta muestra representativa le permite entender cómo la luz natural afecta a las oficinas en todo el complejo. En arquitectura, una muestra es útil para evaluar un aspecto específico, como la luz natural, en un número reducido de oficinas representativas, en lugar de medir cada una. Esto ayuda a obtener una visión general aplicable a todo el edificio.*

1.7.4. Variable

Una característica o atributo que se está midiendo o estudiando en un conjunto de datos. Puede tomar diferentes valores o categorías. Las variables pueden ser cualitativas (categóricas) o cuantitativas (numéricas), y son fundamentales en la recolección y análisis de datos, ya que permiten medir y comparar diferentes aspectos de una población o muestra.

Una variable es como una pieza de información que puede cambiar y que se mide para entender algo más grande. En ingeniería civil y arquitectura, las variables nos ayudan a analizar y mejorar proyectos, ya sea midiendo la cantidad de lluvia que afecta una carretera o la temperatura dentro de un edificio.

Según Rossi y Merenda (2010), las variables se categorizan en cuatro escalas básicas de medición: nominal, ordinal, intervalo y de razón, lo que define cómo se recopilan y analizan los datos en investigaciones científicas y experimentales.

Ejemplos:

1. Considerar que un ingeniero civil está estudiando el impacto de las lluvias en la estabilidad de los taludes en una región montañosa. Aquí, las variables podrían incluir:

- Cantidad de lluvia: medida en milímetros.
- Inclinação del talud: medida en grados.
- Tipo de suelo: por ejemplo, arcilloso, arenoso, rocoso.

Estas variables se miden y analizan para entender cómo cada una afecta la estabilidad de los taludes, ayudando a tomar decisiones para prevenir deslizamientos de tierra.

2. Un arquitecto está diseñando un nuevo edificio y quiere optimizar el confort térmico en los interiores. Las variables en este caso podrían incluir:

- Temperatura interior: medida en grados Celsius.
- Orientación del edificio: norte, sur, este, oeste.
- Material de construcción: concreto, ladrillo, madera.

Estas variables son evaluadas para determinar cuál combinación ofrece el mejor confort térmico durante todo el año.

1.7.5. Dato/Datos

Un dato es un valor individual o una unidad de información que se recoge a partir de una variable y se utilizan para hacer análisis, tomar decisiones y obtener conclusiones en investigaciones y estudios.

Son las observaciones recolectadas (como mediciones, géneros, respuestas de encuesta). Estos datos son la base sobre la cual se construye todo el análisis estadístico y se utilizan para obtener información y tomar decisiones en una variedad de campos.

En ingeniería civil, un dato es una pieza de información específica que se mide o registra. Por ejemplo, si se mide la resistencia del concreto en diferentes muestras, cada medida individual es un dato que ayuda a decidir cuál mezcla es la mejor. Según Xiu (2024), el uso de tecnología de minería de datos en sistemas de ingeniería civil permite extraer, organizar y analizar información crucial como datos de salud estructural en puentes. Esto optimiza los diseños y mejora la calidad de las obras al detectar problemas antes de que se conviertan en fallas graves, mostrando que los datos pueden ser el

núcleo de decisiones críticas en la ingeniería moderna.

En arquitectura, un dato es una información específica que se recoge para analizar un aspecto del diseño. Por ejemplo, si se mide la cantidad de luz que entra por diferentes ventanas, cada medición es un dato que ayuda a decidir qué tipo de ventana es más eficiente. Liu (2021) señala que la integración de datos masivos con tecnología BIM (Building Information Modeling) permite gestionar información arquitectónica detallada como parámetros de edificios y expresiones de información inexactas. Esto no solo mejora la eficiencia del diseño y la operación, sino que también ayuda a solucionar problemas recurrentes en la planificación y mantenimiento de proyectos arquitectónicos.

Ejemplos:

1. Supongamos que un ingeniero civil está llevando a cabo un estudio sobre la resistencia del concreto en diferentes mezclas. Aquí, los datos podrían incluir:

- *Resistencia del concreto en diferentes muestras, por ejemplo, 25 MPa, 30 MPa, 35 MPa.*
- *Tipo de mezcla utilizada, por ejemplo, mezcla A, mezcla B, mezcla C.*

Cada una de estas mediciones proporciona información específica que ayuda al ingeniero a determinar qué mezcla de concreto es la más adecuada para diferentes aplicaciones.

2. Un arquitecto está analizando cómo el diseño de las ventanas afecta la eficiencia energética de un edificio. Los datos podrían ser:

- *Cantidad de luz natural que entra a través de diferentes tipos de ventanas, por ejemplo, 500 lux, 700 lux, 900 lux.*
- *Tipo de ventana utilizada, por ejemplo, ventana de doble acristalamiento, ventana de triple acristalamiento.*

Estos datos ayudan al arquitecto a entender cómo cada tipo de ventana contribuye a la eficiencia energética del edificio.

1.7.6. Parámetro

Un valor numérico que describe una característica de una población,

como la media o la desviación estándar. Es una medida descriptiva que resume o caracteriza una característica de una población completa. Es un valor fijo, aunque a menudo desconocido, que describe una propiedad o atributo de la población, como la media, la desviación estándar o la proporción.

Los parámetros son contrastados con los estadísticos, que son medidas calculadas a partir de una muestra.

Un parámetro es una medida que describe una característica general de una población completa, como el promedio de resistencia de un material o la temperatura media en un edificio. Aunque no siempre se puede medir directamente, se utiliza para comprender y evaluar aspectos importantes en estudios y proyectos. Según Doherty y Hunt (2009), los parámetros también se pueden utilizar en modelos para identificar variables clave que influyen en un fenómeno, lo que permite optimizar la calibración y minimizar errores en los procesos de estimación. Este enfoque es particularmente útil en sistemas complejos como modelos de agua subterránea, donde la precisión en los parámetros es esencial para la gestión de recursos.

Ejemplos:

- 1. Un ingeniero civil está diseñando un puente y necesita conocer la resistencia promedio del concreto utilizado en la construcción. El parámetro en este caso sería la resistencia media del concreto en toda la población de mezclas utilizadas en el proyecto. Aunque el ingeniero puede medir la resistencia en una muestra de mezclas, el parámetro es la resistencia media real para todo el concreto usado en el puente.*
- 2. Se está diseñando un sistema de climatización para un edificio y necesita calcular la temperatura media interior en todo el edificio. El parámetro sería la temperatura media real en todo el edificio durante un período específico. El arquitecto puede recoger datos de una muestra de habitaciones para estimar este parámetro.*

1.7.7. Estadístico

Un valor numérico que describe una característica de una muestra, como la media muestral o la desviación estándar muestral. Es una

medida calculada a partir de una muestra de datos que se utiliza para estimar un parámetro de la población completa. A diferencia del parámetro, que describe una característica de toda la población, el estadístico es una estimación basada en una parte representativa de esa población.

Es útil cuando no se puede medir toda la población, ya que proporciona una forma de hacer inferencias y tomar decisiones basadas en una parte representativa de los datos. Según McCluskey y Lalkhen (2007), los estadísticos permiten realizar estimaciones a partir de datos muestrales, utilizando conceptos de probabilidad para extrapolar información de una muestra hacia una población completa. Este enfoque es crucial en estudios donde medir la población entera sería impráctico o imposible, como el análisis de datos clínicos o sociales.

Ejemplos:

1. *Un ingeniero civil realiza un estudio sobre la resistencia del concreto en un conjunto de edificios nuevos. En lugar de medir el concreto en todos los edificios, toma una muestra de 30 mezclas de concreto. El estadístico en este caso sería la resistencia promedio del concreto calculada a partir de esas 30 muestras.*

Este estadístico se usa para hacer una estimación de la resistencia promedio del concreto en todos los edificios.

En ingeniería civil, un estadístico es un valor calculado a partir de una muestra de datos que se usa para estimar una característica de todo un grupo. Por ejemplo, si medimos la resistencia del concreto en solo algunas mezclas y calculamos el promedio, ese promedio es un estadístico que nos ayuda a inferir la resistencia del concreto en toda la construcción.

2. *Un arquitecto está analizando la eficiencia energética en un edificio. Mide la temperatura en varias habitaciones para evaluar la eficacia del sistema de climatización. El estadístico sería la temperatura promedio medida en esas habitaciones. Este valor se utiliza para estimar cómo funciona el sistema en todo el edificio.*

En arquitectura, un estadístico es una medida calculada a partir de

datos obtenidos de una parte del edificio que se usa para hacer suposiciones sobre el resto. Por ejemplo, si medimos la temperatura en algunas habitaciones y calculamos el promedio, ese promedio es un estadístico que nos ayuda a entender la temperatura en todo el edificio.

1.7.8. Hipótesis

Una afirmación o suposición que se hace sobre una característica de una población o sobre la relación entre variables y que se prueba mediante análisis estadísticos.

Las hipótesis se dividen en dos tipos principales:

- **Hipótesis Nula (H_0):** Afirmación que establece que no hay efecto o diferencia significativa. Según Banerjee et al. (2009), la hipótesis nula se formula como punto de partida para determinar si las evidencias disponibles en los datos son suficientes para rechazarla, basándose en probabilidades estadísticamente definidas.
- **Hipótesis Alternativa (H_1):** Afirmación que propone una diferencia o efecto significativo, que se considera si se rechaza la hipótesis nula. Cramer y Howitt (2004) sostienen que la hipótesis alternativa representa la posición del investigador y se valida únicamente si las pruebas refutan la hipótesis nula con un nivel de confianza establecido.

Una hipótesis es una suposición que se hace para investigar si algo tiene un efecto o diferencia significativa. Se plantea para ser probada con datos y experimentos. Por ejemplo, si se introduce una nueva técnica en un proyecto, la hipótesis puede ser que esta técnica es más efectiva que la anterior. Luego se realizan pruebas para ver si esta suposición es respaldada por los datos.

Ejemplos:

1. *Supongamos que se quiere investigar si un nuevo tipo de mezcla de concreto tiene una resistencia mayor que una mezcla estándar. La hipótesis planteada podría ser:*

- *Hipótesis Nula (H_0): “La resistencia media del nuevo tipo de*

mezcla de concreto es igual a la resistencia media de la mezcla estándar.”

- *Hipótesis Alternativa (H_1): “La resistencia media del nuevo tipo de mezcla de concreto es mayor que la resistencia media de la mezcla estándar.”*

El ingeniero realiza pruebas y recoge datos para determinar si hay suficiente evidencia para rechazar la hipótesis nula y aceptar la hipótesis alternativa.

2. *Se está evaluando si el diseño de ventanas de un nuevo edificio mejora la eficiencia energética en comparación con un diseño tradicional. La hipótesis podría ser:*

- *Hipótesis Nula (H_0): “El diseño de ventanas del nuevo edificio no mejora la eficiencia energética en comparación con el diseño tradicional.”*
- *Hipótesis Alternativa (H_1): “El diseño de ventanas del nuevo edificio mejora la eficiencia energética en comparación con el diseño tradicional.”*

El arquitecto realiza estudios de consumo energético y compara los resultados para probar estas hipótesis.

1.7.9. Prueba de Hipótesis

Un procedimiento estadístico que se utiliza para determinar si una hipótesis es válida o no en función de la evidencia de los datos de una muestra para aceptar o rechazar una hipótesis nula.

La prueba compara los datos observados con lo que se esperaría si la hipótesis nula fuera cierta y se basa en un nivel de significancia predefinido (α), como el 5% o el 1%.

Una prueba de hipótesis es una forma de averiguar si los resultados que obtienes en un experimento o estudio son lo suficientemente diferentes de lo que se esperaría si no hubiera ningún efecto o diferencia real. Se utiliza para tomar decisiones basadas en datos y determinar si lo que estás probando tiene un impacto significativo. Según Sham y Purcell (2014), la prueba de hipótesis es esencial para validar teorías científicas al evaluar probabilidades y establecer límites de significancia en estudios genéticos y científicos, lo que asegura que los resultados no sean producto del azar.

Ejemplos:

1. Se quiere determinar si un nuevo tipo de mezcla de concreto tiene una resistencia superior a una mezcla estándar. El proceso de prueba de hipótesis podría seguir estos pasos:

a) Plantear las Hipótesis:

- Hipótesis Nula (H_0): La resistencia media del nuevo concreto es igual a la resistencia media de la mezcla estándar.
- Hipótesis Alternativa (H_1): La resistencia media del nuevo concreto es mayor que la resistencia media de la mezcla estándar.

b) Recolectar Datos: Medir la resistencia de una muestra de mezclas de concreto nuevas y estándar.

c) Realizar la Prueba Estadística: Usar una prueba t para comparar las medias de las dos muestras.

d) Calcular el Valor P: Determinar la probabilidad de observar los datos actuales si la hipótesis nula fuera cierta.

e) Comparar con el Nivel de Significancia: Si el valor p es menor que el nivel de significancia (por ejemplo, 0.05), se rechaza la hipótesis nula.

f) Conclusión: Si se rechaza la hipótesis nula, se concluye que el nuevo concreto tiene una resistencia significativamente mayor.

2. Se requiere evaluar si el nuevo diseño de ventanas reduce significativamente el consumo de energía en comparación con el diseño tradicional. El proceso de prueba de hipótesis podría ser:

a) Plantear las Hipótesis:

- Hipótesis Nula (H_0): El nuevo diseño de ventanas no reduce el consumo de energía en comparación con el diseño tradicional.
- Hipótesis Alternativa (H_1): El nuevo diseño de ventanas reduce el consumo de energía en comparación con el diseño tradicional.

b) Recolectar Datos: Medir el consumo de energía en edificios con ambos tipos de ventanas.

c) Realizar la Prueba Estadística: Utilizar una prueba t o ANOVA para comparar el consumo de energía entre los dos grupos.

d) Calcular el Valor P: Evaluar la probabilidad de obtener los resultados actuales si la hipótesis nula fuera cierta.

e) Comparar con el Nivel de Significancia: Rechazar la hipótesis nula si el valor p es menor que el nivel de significancia (por ejemplo, 0.05).

f) Conclusión: Si se rechaza la hipótesis nula, se concluye que el nuevo diseño de ventanas es significativamente más eficiente en términos de consumo de energía.

1.7.10. Muestreo Aleatorio

El proceso de selección de una muestra de forma que todos los elementos de la población tienen la misma probabilidad de ser elegido. Este tipo de muestreo es fundamental para garantizar que las muestras sean representativas de la población completa y minimizar sesgos.

El muestreo aleatorio es una técnica para seleccionar una muestra de datos donde cada elemento tiene la misma probabilidad de ser elegido. Esto asegura que la muestra sea representativa de la población completa y ayuda a obtener conclusiones precisas sin tener que examinar todos los elementos. Según Singh (2003), el muestreo aleatorio simple permite la selección de unidades con igual probabilidad en cada paso, lo que lo convierte en uno de los métodos más utilizados y efectivos para asegurar la representatividad de la muestra.

Ejemplos:

1. Para evaluar la calidad del suelo en un área para un proyecto de construcción. En lugar de examinar cada punto del área, el ingeniero usa muestreo aleatorio para seleccionar puntos de prueba. Así:

a) Definir la Población: Todos los puntos de suelo en el área del proyecto.

b) Seleccionar Aleatoriamente: Utilizar un software de muestreo aleatorio para seleccionar una cantidad fija de puntos de prueba (por ejemplo, 20 puntos).

c) Recolectar Datos: Realizar pruebas en los puntos seleccionados para analizar la calidad del suelo.

d) Generalizar Resultados: Usar los datos de la muestra para hacer inferencias sobre la calidad del suelo en toda el área del proyecto. En ingeniería civil, el muestreo aleatorio es como elegir puntos al azar para hacer pruebas en un terreno grande. Así, puedes obtener una

buena idea de la calidad del terreno sin tener que examinar cada punto individualmente.

2. Para evaluar el confort térmico en diferentes oficinas de un edificio. En lugar de medir la temperatura en cada oficina, el arquitecto usa muestreo aleatorio para seleccionar oficinas. Por ejemplo:

*a) **Definir la Población:** Todas las oficinas del edificio.*

*b) **Seleccionar Aleatoriamente:** Elegir al azar un número fijo de oficinas (por ejemplo, 10 oficinas) usando un software o una tabla de números aleatorios.*

*c) **Recolectar Datos:** Medir la temperatura en las oficinas seleccionadas.*

*d) **Generalizar Resultados:** Analizar las mediciones para estimar el confort térmico en el resto del edificio.*

En arquitectura, el muestreo aleatorio es como elegir al azar algunas oficinas en un edificio para medir la temperatura. Esto ayuda a entender cómo es la temperatura en todo el edificio sin tener que medir en todas las oficinas.

1.7.11. Error Estándar

Una medida de la variabilidad o dispersión de un estadístico (como la media) en diferentes muestras tomadas de la misma población. Indica cuánto puede variar el estimador (por ejemplo, la media muestral) de una muestra a otra debido al azar.

El error estándar es una forma de medir cuánto pueden variar los resultados de una muestra respecto a la población completa. Cuanto menor sea el error estándar, más precisas son las estimaciones hechas a partir de la muestra. Es una herramienta clave para entender la confiabilidad de los resultados en investigaciones y estudios. Según Altman y Bland (2005), el error estándar no solo permite evaluar la precisión de las estimaciones al medir la variabilidad de una estadística muestral, sino que también desempeña un papel clave en la construcción de intervalos de confianza, los cuales, al basarse en el error estándar, ofrecen un rango dentro del cual es probable que se ubique el verdadero valor poblacional,

proporcionando así un marco sólido para la inferencia estadística y la toma de decisiones fundamentadas en datos.

Ejemplos:

1. Medir la resistencia de diferentes muestras de concreto para un proyecto. Después de tomar varias muestras, calcula la media de resistencia para cada una. El error estándar de la media de resistencia indica cuánta variabilidad hay en las medias calculadas de diferentes muestras. Un error estándar bajo sugiere que las medias de resistencia de diferentes muestras están bastante cerca unas de otras, mientras que un error estándar alto indica mayor variabilidad.

En ingeniería civil, el error estándar te dice cuánto podrían variar los resultados de las pruebas si tomas diferentes muestras del mismo material. Por ejemplo, si haces varias pruebas de resistencia del concreto, el error estándar te ayuda a entender cuán consistentes son esos resultados.

2. Para evaluar el nivel de confort térmico en diferentes habitaciones de un edificio. Se mide la temperatura en varias habitaciones y calcula la media. El error estándar de la media de temperatura muestra cuánta variabilidad hay entre las temperaturas medidas en las habitaciones. Un error estándar pequeño indica que las temperaturas en las habitaciones son bastante similares, mientras que un error estándar grande sugiere más variabilidad.

En arquitectura, el error estándar muestra cuánta variabilidad hay en las mediciones que haces en diferentes partes de un edificio. Por ejemplo, si mides la temperatura en varias habitaciones, el error estándar te ayuda a entender cuán consistentes son esas mediciones.

1.7.12. Test de Significancia

Un procedimiento estadístico que se utiliza para determinar si los resultados observados son estadísticamente significativos o si podrían haber ocurrido por azar. El objetivo principal de un test de significancia es evaluar si la evidencia en los datos de una muestra es suficiente para rechazar la hipótesis nula en favor de la hipótesis alternativa.

Un test de significancia es una herramienta para tomar decisiones informadas basadas en datos. Permite determinar si los resultados observados son suficientemente improbables bajo una hipótesis inicial para considerar una alternativa. Es fundamental en la investigación para validar si las diferencias o efectos encontrados en los datos son reales y no simplemente el resultado del azar. Según Hayat (2010), el test de significancia es esencial en la investigación para distinguir entre variabilidad natural y efectos reales, mientras que Sham y Purcell (2014) destacan su aplicación en estudios genéticos, donde se manejan múltiples pruebas simultáneas para garantizar la validez de los resultados.

Ejemplos:

- 1. Supongamos que un ingeniero está evaluando si un nuevo tipo de material tiene una resistencia diferente al material estándar. Se realiza un test de significancia para comparar la resistencia media de ambos materiales. Si el valor p calculado es menor que el nivel de significancia elegido (por ejemplo, 0.05), el ingeniero puede concluir que el nuevo material tiene una resistencia significativamente diferente del material estándar.*
- 2. Un arquitecto podría querer determinar si un nuevo diseño de ventanas mejora significativamente la eficiencia energética en comparación con un diseño tradicional. Se realiza un test de significancia para comparar el consumo de energía en edificios con ambos tipos de ventanas. Si el valor p es menor que el nivel de significancia (por ejemplo, 0.01), se concluye que el nuevo diseño de ventanas tiene un efecto significativo en la eficiencia energética.*

1.7.13. Nivel de Significancia

El nivel de significancia (α) es un umbral predefinido en un test estadístico que determina la probabilidad de rechazar incorrectamente la hipótesis nula cuando es verdadera. Representa el riesgo de cometer un error tipo I, es decir, la probabilidad de concluir que existe un efecto o diferencia cuando en realidad no existe.

Los niveles de significancia más comúnmente utilizados son 0.05 (5%) y 0.01 (1%). Un nivel de significancia del 5% implica que hay un 5% de riesgo de rechazar incorrectamente la hipótesis nula, mientras que un nivel de significancia del 1% implica un riesgo del 1%. La selección del nivel de significancia depende del contexto de la investigación y las consecuencias de cometer un error tipo I. En investigaciones donde las consecuencias de un error son graves, se puede optar por un nivel de significancia más bajo. En investigaciones donde las consecuencias de un error son graves, se puede optar por un nivel de significancia más bajo. Según Sham y Purcell (2014), en estudios complejos como los genéticos, donde se realizan múltiples pruebas, es esencial establecer niveles de significancia más estrictos para garantizar la validez de los resultados.

Ejemplos:

1. *Un ingeniero civil está probando la resistencia de un nuevo tipo de material y establece un nivel de significancia del 5%. Si el valor p obtenido es 0.03, que es menor que 0.05, el ingeniero rechazará la hipótesis nula y concluirá que el nuevo material tiene una resistencia significativamente diferente.*
2. *Un arquitecto evalúa la eficiencia energética de un nuevo diseño de ventanas y elige un nivel de significancia del 1%. Si el valor p es 0.007, que es menor que 0.01, se rechazará la hipótesis nula y se concluirá que el nuevo diseño mejora significativamente la eficiencia energética.*

1.7.14. Error Tipo I

Ocurre cuando se rechaza incorrectamente una hipótesis nula (H_0) verdadera en una prueba de hipótesis. Este error también se conoce como *falsos positivos* y está asociado con el nivel de significancia (α) de una prueba estadística. El nivel de significancia representa la probabilidad de cometer un error tipo I. Según Efsthathiou (2006), los errores tipo I son comunes en pruebas de significancia, especialmente cuando las distribuciones de datos no están

adecuadamente modeladas, lo que subraya la importancia de seleccionar correctamente los niveles de significancia y las herramientas estadísticas para minimizar este tipo de error.

Ejemplos:

1. Supongamos que se está evaluando si un nuevo material de construcción tiene una resistencia mayor que el material estándar. Un error tipo I ocurriría si el ingeniero concluye que el nuevo material es más resistente cuando en realidad no hay una diferencia significativa.
2. Se está evaluando si un nuevo tipo de aislamiento térmico para ventanas reduce significativamente el consumo de energía en comparación con un aislamiento estándar. Un error tipo I en este contexto ocurriría si, en realidad, el nuevo aislamiento no tiene un impacto significativo en la reducción del consumo de energía (es decir, la hipótesis nula (H_0) es verdadera), pero el análisis estadístico concluye incorrectamente que sí lo tiene. En otras palabras, el arquitecto podría creer erróneamente que el nuevo aislamiento es más eficiente cuando en realidad no hay una diferencia real.

1.7.15. Error Tipo II

Un error tipo II ocurre cuando no se rechaza la hipótesis nula (H_0) cuando en realidad la hipótesis alternativa (H_1) es verdadera. Este error también se conoce como *falsos negativos*. El riesgo de cometer un error tipo II está asociado con el poder de la prueba, que es la probabilidad de detectar un efecto real cuando existe. Según Jiménez-Gamero y Analla (2023), los errores tipo II tienen implicaciones importantes, ya que no reconocer diferencias verdaderas puede limitar avances científicos y retrasar la implementación de soluciones efectivas, especialmente en investigaciones críticas como salud y toxicología.

Ejemplos:

1. Imaginemos que un ingeniero está evaluando si un nuevo material de construcción tiene una resistencia superior a la del material

estándar. Un error tipo II ocurre si el ingeniero no detecta una diferencia significativa en la resistencia cuando en realidad el nuevo material es más resistente.

- 2. Un arquitecto está evaluando si un nuevo tipo de iluminación LED reduce significativamente el consumo de energía en comparación con la iluminación convencional. La hipótesis nula es que el nuevo tipo de iluminación no tiene un impacto significativo en el consumo de energía, mientras que la hipótesis alternativa es que sí lo tiene. Un error tipo II en este contexto ocurriría si, en realidad, la iluminación LED sí reduce significativamente el consumo de energía (la hipótesis alternativa es verdadera), pero el análisis no detecta esta diferencia y no rechaza la hipótesis nula. En otras palabras, el arquitecto podría creer incorrectamente que la iluminación LED no es más eficiente cuando, en realidad, sí lo es.*

1.7.16. Diseño Experimental

El diseño experimental implica la planificación cuidadosa de cómo se recopilan los datos en un estudio, considerando aspectos como la asignación de grupos, la selección de muestras y el control de variables de confusión.

Este enfoque sistemático permite estructurar, ejecutar y analizar experimentos con el objetivo de investigar y validar hipótesis de manera eficiente. Al manipular variables independientes para observar su impacto en variables dependientes, el diseño experimental resulta esencial para establecer relaciones causales y minimizar la influencia de factores externos que podrían distorsionar los resultados.

En la investigación científica y técnica, el diseño experimental desempeña un papel crucial para investigar causalidades y evaluar la efectividad de diversas intervenciones o tratamientos. Una planificación adecuada y una ejecución rigurosa aseguran resultados válidos y confiables, facilitando decisiones prácticas en campos como la ingeniería civil y la arquitectura.

Hall (2011) resalta que, más allá de las ciencias naturales, este enfoque también ha sido adaptado a disciplinas como el diseño industrial, donde modelos no lineales y perspectivas

interdisciplinarias han optimizado procesos creativos y productivos.

Ejemplos:

1. *Un ingeniero civil quiere evaluar el impacto de diferentes aditivos en la resistencia del concreto.*

a) *Planteamiento del Experimento:*

Tipo de aditivo utilizado.

Resistencia del concreto medido en Megapascals (MPa).

Proporciones de los otros ingredientes del concreto, condiciones de curado, temperatura, etc.

b) *Grupos:*

Muestras de concreto con diferentes tipos de aditivos.

Muestras de concreto sin aditivo (o con el aditivo estándar).

c) *Aleatorización: Se asignan aleatoriamente diferentes aditivos a las muestras para asegurar que no haya sesgos en la asignación.*

d) *Muestreo: Selección aleatoria de muestras de concreto para asegurar representatividad.*

e) *Diseño de Experimento: Se evalúa el efecto de varios aditivos y sus combinaciones sobre la resistencia del concreto.*

2. *Un arquitecto quiere comparar la eficiencia energética de diferentes tipos de ventanas en un edificio.*

a) *Planteamiento del Experimento:*

Tipo de ventana (por ejemplo, ventanas de doble acristalamiento, ventanas de triple acristalamiento, ventanas con recubrimiento especial).

Consumo de energía medido en kilovatios hora (kWh).

Tamaño de las ventanas, orientación de las habitaciones, condiciones climáticas durante el período de prueba.

b) *Grupos*

Habitaciones equipadas con diferentes tipos de ventanas.

Habitaciones con ventanas estándar para comparar el consumo de energía.

c) *Aleatorización: Asignación aleatoria de diferentes tipos de ventanas a las habitaciones para evitar sesgos.*

d) *Muestreo: Selección de habitaciones en el edificio de manera que representen adecuadamente las condiciones generales.*

e) *Diseño de Experimento: Se mide el consumo de energía en cada habitación durante diferentes estaciones para evaluar el rendimiento a lo largo del tiempo.*

1.8. Tipos de datos

Los datos estadísticos pueden ser cualitativos o cuantitativos.

1.8.1. Datos cualitativos

Los datos cualitativos representan categorías o atributos que se dividen en diferentes grupos según características no numéricas. Estos datos son esenciales para describir cualidades o características de fenómenos u objetos que no pueden ser medidos numéricamente. En investigaciones científicas y sociales, los datos cualitativos permiten explorar y entender fenómenos desde una perspectiva interpretativa. Según Seers (2011), estos datos suelen expresarse en palabras o textos, y se recopilan a través de entrevistas, observaciones o documentos. Este enfoque permite comprender significados, experiencias y contextos, proporcionando una riqueza de detalles que complementa los datos cuantitativos.

Ejemplos:

- *Tipo de Material de Construcción*
- *Nivel de Riesgo Sísmico*
- *Categorización de Impacto Ambiental*
- *Estado de Conservación de una Estructura*
- *Niveles de Cumplimiento Normativo*
- *Priorización de Riesgos*

1.8.2. Clasificación datos cualitativos

Los datos cualitativos se pueden clasificar en dos categorías principales en función de su naturaleza y propósito. Estas categorías son:

1.8.2.1. Datos Cualitativos Nominales

Los datos nominales son datos que se utilizan para categorizar o

clasificar elementos en grupos o categorías. Las categorías en los datos nominales no tienen un orden inherente y no se pueden cuantificar numéricamente.

Ejemplos:

- **Clasificación de Suelo:** Los tipos de suelo en un sitio de construcción, como arcilla, arena, limo o roca, se clasifican nominalmente en función de su composición.
- **Categoría de Uso de Suelo:** En la planificación urbana y en proyectos de ingeniería civil, las áreas de tierra se clasifican nominalmente según su uso, como residencial, comercial, industrial o recreativo.
- **Tipo de Puente:** En ingeniería de puentes, se pueden clasificar los puentes según su diseño estructural, como puentes colgantes, puentes de arco o puentes de vigas.
- **Categoría de Riesgo Sísmico:** En áreas propensas a terremotos, los lugares se pueden categorizar nominalmente en función de su nivel de riesgo sísmico, como zona de alto riesgo o zona de riesgo moderado.
- **Estilo Arquitectónico:** Los edificios pueden ser clasificados nominalmente según su estilo arquitectónico, como gótico, renacentista, moderno, art déco, etc.
- **Color de Fachada:** El color de la fachada de un edificio se clasifica nominalmente según su tono, como rojo, verde, azul, blanco, etc.
- **Tipo de Material de Revestimiento:** Los materiales utilizados en la fachada de un edificio, como ladrillo, piedra natural, estuco; se clasifican nominalmente según su tipo.
- **Categoría de Uso del Espacio:** Los espacios en un edificio se pueden clasificar nominalmente según su función, como área de recepción, oficina, sala de reuniones, etc.
- **Tipo de Ventana:** Las ventanas en un diseño arquitectónico se pueden clasificar nominalmente según su tipo, como ventana panorámica, ventana de guillotina, ventana de arco, etc.

1.8.2.2. Datos Cualitativos Ordinales

Los datos ordinales son datos que se utilizan para clasificar elementos en categorías con un orden jerárquico o relativo. Aunque

las categorías en datos ordinales tienen un orden, la distancia entre las categorías no es constante ni cuantificable.

Ejemplos:

- **Nivel de Severidad de Daño Estructural:** Después de un terremoto, se pueden asignar niveles de severidad de daño estructural en una escala ordinal, como leve, moderado, grave y colapso.
- **Clasificación de Riesgo de Inundación:** En proyectos de ingeniería hidráulica, se utiliza una escala ordinal para clasificar áreas según su riesgo de inundación, como bajo, moderado, alto y muy alto.
- **Calificación de Calidad del Aire:** En la evaluación ambiental de proyectos, se puede utilizar una escala ordinal para calificar la calidad del aire en términos de buena, moderada, mala y muy mala.
- **Grado de Compactación del Suelo:** Al evaluar la calidad del suelo para la construcción, se utiliza una escala ordinal para clasificar el grado de compactación, como baja, media, alta y óptima.
- **Categorización de Capacidad de Carga de Puentes:** Los ingenieros pueden asignar categorías ordinales de capacidad de carga a puentes, como capacidad sobresaliente, capacidad adecuada, capacidad reducida y capacidad crítica.
- **Nivel de Confort Térmico:** En el diseño de edificios, se utiliza una escala ordinal para calificar el nivel de confort térmico de espacios interiores, como muy cómodo, cómodo, marginalmente cómodo e incómodo.
- **Grado de Iluminación Natural:** En arquitectura, se puede calificar la cantidad de iluminación natural en una habitación en términos de muy luminosa, luminosa, poco luminosa y oscura.
- **Categorización de Eficiencia Energética:** Los edificios se pueden clasificar en categorías ordinales de eficiencia energética, como muy eficiente, eficiente, moderadamente eficiente e ineficiente.
- **Nivel de Privacidad:** En el diseño de espacios interiores, se utiliza una escala ordinal para calificar el nivel de privacidad de áreas como privado, semiprivado, semipúblico y público.
- **Calificación de Accesibilidad:** En la planificación de edificios y espacios públicos, se utiliza una escala ordinal para clasificar la

accesibilidad, como totalmente accesible, accesible con limitaciones, poco accesible y no accesible.

Estas dos categorías principales, datos cualitativos nominales y datos cualitativos ordinales, se utilizan para describir y clasificar atributos, características o cualidades de elementos, pero no se expresan en términos de cantidades numéricas. Los datos cualitativos son esenciales en muchas disciplinas, como sociología, psicología, ciencias sociales, diseño y más, donde la información no siempre se puede medir de manera numérica, pero aún es importante para la toma de decisiones y el análisis.

1.8.3. Datos cuantitativos

Son valores numéricos que pueden ser medidos y comparados. Consisten en números que representan conteos o mediciones, y permiten expresar medidas, cantidades o atributos que pueden ser cuantificados. Estos datos se someten a operaciones matemáticas y estadísticas para obtener conclusiones significativas.

Los datos cuantitativos son fundamentales en el análisis estadístico, ya que proporcionan una forma estructurada de interpretar fenómenos y relaciones. Según Albers (2017), el análisis de datos cuantitativos implica no solo *crunching numbers* ('hacer cálculos intensivos'), sino el uso de técnicas estadísticas para descubrir patrones y relaciones profundas en los datos, lo que permite tomar decisiones informadas y fundamentadas.

Ejemplos:

- *Longitud de una Viga*
- *Diámetro de un Tubo de Drenaje*
- *Carga de Viento*
- *Presión del Suelo*
- *Flujo de Agua en un Canal*
- *Peso de Materiales de Construcción*
- *Nivel de Vibración en una Estructura*
- *Costo del Proyecto*

- *Densidad del Suelo.*
- *Tiempo de Construcción*

1.8.4. Clasificación datos cuantitativos

Los datos cuantitativos se pueden clasificar en dos categorías principales según su naturaleza y cómo se pueden medir. Estas categorías son:

1.8.4.1. Datos Cuantitativos Discretos:

Estos datos representan valores que son contables y se expresan generalmente como números enteros. Los valores no pueden tomar infinitas posibilidades dentro de un rango, sino que se limitan a valores específicos.

Ejemplos:

- *Cantidad de Vigas en una Estructura de Puente.*
- *Número de Pilotes en una Cimentación de Edificio*
- *Cantidad de Luminarias en un Sistema de Iluminación Urbana*
- *Cantidad de Tuberías en un Sistema de Alcantarillado*
- *Número de Habitaciones en un Hotel*
- *Cantidad de Asientos en un Teatro*
- *Número de Pisos en un Edificio de Oficinas*
- *Cantidad de Escalones en una Escalera*
- *Número de Puertas en un Diseño de Vivienda*

1.8.4.2. Datos Cuantitativos Continuos

Estos datos representan valores que pueden tomar un rango infinito de posibilidades dentro de un intervalo. Los valores se pueden medir con precisión hasta un cierto número decimal.

Ejemplos:

- *Espesor de una Pared*
- *Longitud de una Carretera*
- *Altura de una Presa*
- *Temperatura del Hormigón*

- *Flujo de Agua en un Río*
- *Altura de un Techo*
- *Área de Ventanas en una Fachada*
- *Volumen de una Habitación*
- *Anchura de una Escalera*
- *Área de Superficie de una Fachada*

Esta clasificación se basa en la naturaleza de los datos y cómo se pueden medir. Los datos cuantitativos discretos se caracterizan por ser contables y tener valores distintos, mientras que los datos cuantitativos continuos pueden tomar un rango infinito de valores dentro de un intervalo y se pueden medir con precisión hasta cierto número de decimales. Esta distinción es relevante en el análisis estadístico y en la elección de las técnicas estadísticas adecuadas para analizar estos tipos de datos.

1.9. Tipos de medición

Los niveles de medición, también conocidos como escalas de medición, se utilizan para describir la naturaleza de los datos y determinar qué operaciones matemáticas y estadísticas se pueden aplicar a esos datos.

Los cuatro niveles de medición son:

1.9.1. Nivel de Medición Nominal

En este nivel, los datos se clasifican en categorías o grupos. Las categorías no tienen un orden intrínseco y no se pueden someter a operaciones matemáticas.

Ejemplos:

- *Categorización de Puentes: Los puentes pueden clasificarse nominalmente en función de su diseño estructural, como puente colgante, puente de arco, puente de vigas o puente atirantado.*
- *Tipo de Tuberías de Agua: En proyectos de abastecimiento de agua, las tuberías se pueden clasificar nominalmente según su material, como tuberías de PVC, tuberías de hierro fundido o*

tuberías de acero.

- *Estilo Arquitectónico:* Los edificios se pueden clasificar nominalmente según su estilo arquitectónico, como gótico, renacentista, moderno, art déco o minimalista.
- *Categorización de Materiales de Revestimiento:* Los materiales utilizados en la fachada de un edificio se pueden clasificar nominalmente según su tipo, como ladrillo, piedra natural, estuco o paneles de vidrio.

1.9.2. Nivel de Medición Ordinal

En este nivel, los datos se clasifican en categorías con un orden relativo. Las categorías tienen un orden, pero la distancia entre ellas no es constante ni cuantificable. Se pueden realizar operaciones de orden, como comparaciones y clasificaciones.

Ejemplos:

- **Clasificación de Riesgo Sísmico:** Las zonas sísmicas se pueden clasificar en términos de su riesgo sísmico en una escala ordinal, como zona de alto riesgo, zona de riesgo moderado y zona de bajo riesgo.
- **Nivel de Daño Estructural:** Después de un terremoto, los niveles de daño estructural se pueden clasificar ordinalmente en función de su gravedad, como leve, moderado, grave y colapso.
- **Categorización de Capacidad de Carga de Puentes:** Los puentes pueden ser categorizados en términos de su capacidad de carga en una escala ordinal, como capacidad sobresaliente, capacidad adecuada, capacidad reducida y capacidad crítica.
- **Grado de Compactación del Suelo:** El grado de compactación del suelo se puede clasificar en una escala ordinal, como baja, media, alta y óptima.
- **Nivel de Inundación:** En proyectos de ingeniería hidráulica, las áreas se pueden clasificar ordinalmente en términos de su nivel de inundación, como ninguna inundación, inundación leve, inundación moderada e inundación grave.
- **Nivel de Confort Térmico:** En el diseño de edificios, los niveles de confort térmico se pueden clasificar ordinalmente en términos de muy cómodo, cómodo, marginalmente cómodo e incómodo.

- **Categoría de Eficiencia Energética:** Los edificios pueden ser categorizados en términos de su eficiencia energética en una escala ordinal, como muy eficiente, eficiente, moderadamente eficiente y ineficiente.
- **Grado de Iluminación Natural:** La cantidad de iluminación natural en un espacio interior se puede clasificar ordinalmente en términos de muy luminoso, luminoso, poco luminoso y oscuro.
- **Calificación de Accesibilidad:** Los espacios públicos y edificios se pueden clasificar en términos de su accesibilidad en una escala ordinal, como totalmente accesible, accesible con limitaciones, poco accesible y no accesible.
- **Categorización de Privacidad:** Los espacios en un edificio se pueden clasificar ordinalmente en términos de su nivel de privacidad, como privado, semiprivado, semipúblico y público.

1.9.3. Nivel de Medición de Intervalo

En este nivel, los datos se clasifican en categorías con un orden y una distancia constante y cuantificable entre las categorías. No hay un punto de partida absoluto (cero absolutos) en esta escala. Se pueden realizar operaciones de orden y cálculos de diferencia.

Ejemplos:

- **Temperatura en Grados Celsius:** La temperatura en grados Celsius se mide en una escala de intervalo donde la distancia entre cada grado es constante. Se utiliza en la ingeniería civil para evaluar el comportamiento térmico de estructuras y materiales.
- **Índice de Calidad del Aire (ICA):** El Índice de Calidad del Aire es una medida que utiliza una escala de intervalo para evaluar la calidad del aire en función de la concentración de contaminantes atmosféricos. Se aplica en proyectos de control de la contaminación del aire.
- **Densidad de Tráfico en Carreteras:** La densidad de tráfico se mide en vehículos por kilómetro y representa una escala de intervalo utilizada para analizar la congestión vial y la capacidad de las carreteras.
- **Nivel de Radiación:** En proyectos de ingeniería nuclear, el nivel de radiación se mide en unidades de exposición y representa una

escala de intervalo que evalúa la seguridad radiológica.

- **Índice de Calidad del Agua (ICA):** El Índice de Calidad del Agua se utiliza en la evaluación de la calidad del agua en ríos y cuerpos de agua. Se basa en una escala de intervalo que considera parámetros como la concentración de contaminantes.
- **Temperatura Interior:** La temperatura dentro de un edificio se mide en grados Celsius o Fahrenheit, utilizando una escala de intervalo para controlar el confort térmico de los ocupantes.
- **Humedad Relativa:** La humedad relativa del aire en espacios interiores se mide en porcentaje y se utiliza en el diseño de sistemas de climatización y control ambiental.
- **Nivel de Iluminación Artificial:** La intensidad de la iluminación artificial se mide en lux o lúmenes por metro cuadrado, lo que representa una escala de intervalo utilizada en el diseño de iluminación de interiores.
- **Espesor de Aislamiento Térmico:** El espesor del aislamiento térmico en una estructura se mide en milímetros o centímetros y utiliza una escala de intervalo para controlar la eficiencia energética.
- **Distancia entre Columnas:** En el diseño arquitectónico, la distancia entre columnas en un espacio se mide en metros o pies, utilizando una escala de intervalo para distribuir la carga estructural de manera adecuada.

1.9.4. Nivel de Medición de Razón

Este es el nivel más alto de medición, que incluye todas las características de los niveles anteriores. Los datos se clasifican en categorías con un orden, distancia constante y un punto de partida absoluto (cero absolutos). Se pueden realizar operaciones de orden, cálculos de diferencia, multiplicación y división.

Ejemplos:

- **Longitud de un Puente en Metros:** La longitud de un puente se mide en metros, y esta medida tiene un punto de partida absoluto en el cero, lo que permite realizar operaciones de adición y sustracción. Por ejemplo, un puente de 100 metros es el doble de largo que un puente de 50 metros.

- **Peso de Materiales de Construcción en Kilogramos:** El peso de los materiales de construcción se mide en kilogramos y tiene unos cero absolutos, lo que permite realizar operaciones de suma y resta. Por ejemplo, la suma de 1000 kg de concreto y 500 kg de acero es igual a 1500 kg.
- **Área de Terreno en Metros Cuadrados:** El área de un terreno se mide en metros cuadrados, y se pueden realizar operaciones de multiplicación y división. Por ejemplo, un terreno de 500 m² es el doble de grande que un terreno de 250 m².
- **Presión de Agua en Pascal (Pa):** La presión del agua se mide en pascal (Pa), y esta medida tiene unos cero absolutos. Permite realizar operaciones matemáticas, como sumar o restar presiones en diferentes puntos de una tubería.
- **Tiempo en Segundos:** El tiempo se mide en segundos y tiene unos cero absolutos, lo que permite realizar operaciones matemáticas como sumar, restar, multiplicar y dividir. Por ejemplo, 60 segundos son equivalentes a un minuto.
- **Altura de un Edificio en Metros:** La altura de un edificio se mide en metros y tiene unos cero absolutos, lo que permite realizar operaciones matemáticas para comparar alturas de diferentes edificios.
- **Volumen de una Habitación en Metros Cúbicos:** El volumen de una habitación se mide en metros cúbicos y permite realizar operaciones matemáticas de multiplicación y división para calcular la capacidad y la ventilación requerida.
- **Peso de una Estructura en Toneladas:** El peso de una estructura se mide en toneladas métricas, y esta medida tiene unos cero absolutos, lo que facilita el cálculo de la carga y la capacidad de soporte.
- **Presión del Aire en una Habitación en Pascal (Pa):** La presión del aire en una habitación se mide en pascal (Pa), lo que permite realizar operaciones matemáticas para calcular cambios de presión en sistemas de climatización.
- **Distancia Entre Pilares en Metros:** La distancia entre pilares se mide en metros y permite realizar operaciones matemáticas para determinar la distribución de cargas y la estabilidad estructural.

Comprender el nivel de medición de los datos, ya que determina las operaciones estadísticas y matemáticas adecuadas que se pueden aplicar a esos datos. Los niveles de medición van desde los datos más simples y cualitativos (nominales) hasta los más complejos y cuantitativos (de razón).

1.10. El proceso de análisis estadístico

El proceso de análisis estadístico consiste en una serie de pasos sistemáticos diseñados para examinar y extraer conclusiones a partir de datos, con el objetivo de responder preguntas específicas de investigación y tomar decisiones informadas.

1.10.1. Recopilación de Datos

Esta etapa inicial implica la obtención de datos relevantes para el problema o la pregunta de investigación. Los datos pueden recopilarse mediante encuestas, observaciones, experimentos o fuentes secundarias. La calidad de los datos recopilados es fundamental, ya que errores o sesgos en esta etapa pueden comprometer los resultados. Jakobsen et al. (2014) destacan la importancia de garantizar la validez de los datos para obtener conclusiones fiables.

1.10.2. Organización de Datos

Una vez recopilados, los datos deben organizarse adecuadamente, lo que incluye ingresar información en hojas de cálculo o software estadístico, crear tablas y categorizar variables de interés. Este paso asegura un manejo estructurado de la información y permite identificar patrones preliminares. Thulasingam y Premarajan (2018) señalan que una buena organización de datos facilita su análisis posterior y ayuda a reducir errores en las etapas siguientes.

1.10.3. Limpieza de Datos

La limpieza de datos es fundamental para eliminar errores, valores atípicos y datos faltantes que puedan afectar los resultados. Este proceso incluye verificar la consistencia de las variables y aplicar

métodos específicos para manejar datos ausentes. Fife (2019) enfatiza que este paso es fundamental para garantizar la integridad de los resultados y evitar interpretaciones erróneas.

1.10.4. Exploración de Datos

La exploración de datos implica un análisis preliminar mediante técnicas gráficas y estadísticas descriptivas. Este paso permite identificar patrones, anomalías y características importantes antes de realizar análisis más complejos. Kamin (2010) sugiere que las herramientas gráficas son especialmente útiles para obtener una visión inicial de los datos y detectar posibles inconsistencias.

1.10.5. Análisis Estadístico Descriptivo

En esta etapa se calculan medidas como la media, mediana, moda, desviación estándar y otros indicadores para resumir y describir los datos en detalle. Estas estadísticas proporcionan una visión general de la distribución y variabilidad de los datos. Hahn et al. (2019) destacan que estas medidas son esenciales para establecer una base sólida antes de realizar análisis inferenciales.

1.10.6. Análisis Estadístico Inferencial

Esta etapa se centra en extraer conclusiones sobre una población más amplia basándose en los datos de una muestra. Incluye pruebas de hipótesis, análisis de varianza y regresión. Trafimow et al. (2024) argumentan que los métodos inferenciales son fundamentales para evaluar relaciones y efectos con rigor estadístico, permitiendo tomar decisiones basadas en evidencias.

1.10.7. Interpretación de resultados

Los resultados obtenidos se interpretan en el contexto del problema o pregunta de investigación. Este paso implica evaluar la relevancia práctica y estadística de los hallazgos. Doncaster y Ezard (2017) resaltan la importancia de interpretar los resultados en relación con el diseño del estudio para evitar confundir significancia estadística con relevancia práctica.

1.10.8. Presentación de Resultados

La comunicación clara de los resultados mediante gráficos, tablas e informes es esencial para que los hallazgos sean comprensibles para el público objetivo. Thulasingham y Premarajan (2018) señalan que una presentación efectiva mejora la utilidad de los datos al destacar patrones relevantes y apoyar decisiones basadas en ellos.

1.10.9. Toma de Decisiones

Finalmente, los resultados obtenidos se utilizan para guiar decisiones prácticas en negocios, políticas públicas o investigación. Storey y Tibshirani (2003) enfatizan que el uso adecuado de análisis estadísticos en la toma de decisiones mejora la precisión y confiabilidad de las acciones basadas en datos.

1.11. Ejemplos de aplicación del proceso de análisis estadístico

Se presentan guías de trabajo que ilustran la relevancia de las estadísticas en la ingeniería civil. Desde la evaluación de riesgos en la construcción de presas hasta la programación eficiente de proyectos de transporte, estos ejemplos destacan cómo las estadísticas informan y mejoran la toma de decisiones en este campo crítico.

El objetivo de las guías desarrolladas que describen los pasos del análisis estadístico en forma general es proporcionar un marco estructurado y sistemático para realizar un análisis de datos riguroso y efectivo. En el contexto específico de evaluar riesgos en infraestructuras, como una represa hidroeléctrica, estos pasos ayudan a: garantizar la calidad de los datos, facilitar la comprensión de datos complejos, permitir la evaluación de riesgos y probabilidades, soportar la toma de decisiones informadas, mejorar la comunicación de resultados y optimizar la gestión de riesgos.

1.11.1. Guía de trabajo 1: Evaluación de Riesgos Sísmicos en un Puente de Importancia Crítica

En la ingeniería civil, la seguridad de las infraestructuras es una prioridad fundamental. Un caso emblemático que ilustra la relevancia de las estadísticas en este campo es la evaluación de riesgos sísmicos en un puente de importancia crítica ubicado en una región propensa a terremotos. Este caso de estudio demuestra cómo las estadísticas desempeñan un papel esencial en la toma de decisiones informadas y en la ingeniería sísmica.

El puente en cuestión es una vía de transporte que conecta dos áreas urbanas densamente pobladas. Dada su importancia estratégica, la seguridad y la capacidad de resistencia sísmica son esenciales para garantizar la continuidad de las operaciones de transporte y, sobre todo, la seguridad de quienes lo utilizan.

Paso 1: Recopilación de Datos

- **Datos sísmicos:** Registros históricos de actividad sísmica (magnitud, frecuencia, duración).
- **Datos estructurales:** Información sobre el diseño del puente, materiales, características de construcción.
- **Datos de carga:** Información sobre el tráfico (volumen, tipo de vehículos, peso).

Paso 2: Organización de Datos

- **Clasificación:** Agrupar datos sísmicos por magnitud y frecuencia.
- **Codificación:** Asignar códigos a diferentes tipos de datos estructurales y de carga.
- **Estructuración:** Crear bases de datos y tablas que faciliten el análisis.

Paso 3: Limpieza de Datos

- **Verificación de precisión:** Corregir errores e inconsistencias en los datos sísmicos y estructurales.

- **Eliminación de duplicados:** Eliminar registros redundantes o repetidos.
- **Tratamiento de datos faltantes:** Imputar valores faltantes o eliminar registros incompletos si es necesario.

Paso 4: Simulación Sísmica

- **Se realizan simulaciones sísmicas** basadas en los modelos de amenazas para evaluar cómo el puente resistirá diferentes escenarios de terremotos. Esto incluye el análisis de tensiones y deformaciones en los elementos estructurales.
- **Visualización inicial:** Crear gráficos y diagramas para entender la distribución de los datos sísmicos y estructurales.
- **Identificación de patrones:** Examinar tendencias y relaciones preliminares entre los datos.

Paso 5: Evaluación de Riesgos y Toma de Decisiones

- **Cálculo de medidas de tendencia central:** Media, mediana, y moda de la magnitud de los terremotos.
- **Cálculo de medidas de dispersión:** Desviación estándar, varianza, y rango para datos sísmicos y de carga.
- **Distribución de frecuencias:** Analizar la frecuencia de diferentes niveles de magnitud sísmica y tipos de cargas en el puente.

Paso 6: Diseño y Construcción Resistentes a Sismos

- **Estimación de parámetros:** Utilizar técnicas de inferencia para estimar la probabilidad de fallos estructurales.
- **Pruebas de hipótesis:** Evaluar hipótesis sobre la resistencia del puente frente a distintos niveles de actividad sísmica.
- **Modelado predictivo:** Aplicar modelos estadísticos para prever el comportamiento del puente en diferentes escenarios sísmicos.

Paso 7: Monitoreo Continuo

- **Interpretar estimaciones y pruebas de hipótesis:** Evaluar si los resultados cumplen con los estándares de seguridad y los

requisitos de diseño.

- **Analizar implicaciones prácticas:** Evaluar cómo los resultados afectan la seguridad del puente y las decisiones de mantenimiento.

Paso 8: Presentación de Resultados

- **Preparar informes detallados:** Incluir gráficos, tablas y explicaciones sobre los hallazgos.
- **Comunicar a partes interesadas:** Presentar los resultados a ingenieros, autoridades de transporte y otras partes interesadas.

Paso 9. Toma de Decisiones

- **Desarrollar recomendaciones:** Proponer medidas de mitigación y mejoras estructurales basadas en los resultados.
- **Planificación de acciones:** Crear planes para reforzar el puente y asegurar su continuidad operativa en caso de actividad sísmica.

1.11.2. Guía de trabajo 2: Evaluación de Riesgos en una Represa Hidroeléctrica

El uso de estadísticas para evaluar la probabilidad de inundación y el riesgo de colapso de una represa en función de datos hidrológicos históricos y modelos de precipitación.

En la ingeniería civil, la seguridad de las infraestructuras es una prioridad fundamental.

Paso 1: Recopilación de Datos

Se recopilan datos:

- **Datos hidrológicos históricos:** Información sobre caudales de ríos, niveles de agua en la represa, y eventos de inundación pasados.
- **Datos de precipitación:** Registros históricos de precipitación en la cuenca hidrográfica, incluyendo intensidad y duración de tormentas.
- **Datos estructurales de la represa:** Características del diseño, capacidad de almacenamiento, y medidas de seguridad.

Paso 2: Organización de Datos

- **Clasificación de datos hidrológicos:** Agrupar datos por períodos de tiempo, tipo de evento (lluvias extremas, caudales máximos).
- **Codificación de precipitación:** Registrar datos de precipitación por intensidad y duración.
- **Estructuración de bases de datos:** Crear tablas para datos hidrológicos, de precipitación y características estructurales de la represa.

Paso 3: Limpieza de Datos

- **Verificación de precisión:** Corregir errores en los registros de precipitación y caudales de agua.
- **Eliminación de duplicados:** Asegurarse de que no haya registros repetidos en los datos.

- **Tratamiento de datos faltantes:** Imputar valores faltantes o eliminar registros incompletos con métodos estadísticos adecuados.

Paso 4: Exploración de Datos

- **Visualización inicial:** Crear gráficos de series temporales para los datos de precipitación y caudales de agua.
- **Identificación de patrones:** Analizar la relación entre eventos de precipitación intensa y caudales máximos en la represa.

Paso 5: Análisis Estadístico Descriptivo

- **Cálculo de medidas de tendencia central:** Media, mediana y moda de la precipitación y caudales de agua.
- **Cálculo de medidas de dispersión:** Desviación estándar y varianza para evaluar la variabilidad en los datos.
- **Distribución de frecuencias:** Analizar la frecuencia de eventos extremos de precipitación y caudales.

Paso 6: Análisis Estadístico Inferencial

- **Modelos de frecuencia de eventos:** Utilizar distribuciones estadísticas (como la distribución de Gumbel) para estimar la probabilidad de eventos extremos de precipitación y caudales.
- **Modelos hidrológicos:** Aplicar modelos que relacionen la precipitación con el caudal de salida de la represa.
- **Pruebas de hipótesis:** Evaluar hipótesis sobre la capacidad de la represa para manejar diferentes magnitudes de caudales y niveles de precipitación.

Paso 7: Interpretación de Resultados

- **Evaluación de probabilidades de inundación:** Interpretar la probabilidad de que eventos de precipitación intensa resulten en niveles de agua que superen la capacidad de la represa.
- **Análisis de riesgo de colapso:** Determinar el riesgo de colapso basado en la capacidad de la represa frente a eventos extremos de precipitación y caudales.

8. Presentación de Resultados

- **Preparar informes detallados:** Incluir gráficos, tablas y descripciones de los hallazgos del análisis de riesgo de inundación y colapso.
- **Comunicar a partes interesadas:** Presentar los resultados a ingenieros de la represa, autoridades de regulación y equipos de gestión de emergencias.

9. Toma de Decisiones

- **Proponer medidas de gestión de riesgos,** como mejoras en la capacidad de almacenamiento de la represa, sistemas de alerta temprana, y protocolos de emergencia.
- **Crear planes de contingencia y mantenimiento** para reducir el riesgo de inundación y colapso.

Ejercicios a desarrollar del Capítulo 1

Sesión 1.1 Identifique en los siguientes enunciados los conceptos de Población

1. Un arquitecto quiere estudiar las características de todas las casas construidas en un nuevo desarrollo residencial. ¿Cuál es la población en este estudio?
2. Un ingeniero civil quiere analizar la resistencia de un tipo específico de concreto. Recolecta muestras de 10 construcciones diferentes que usaron ese concreto. ¿Cuál es la población en este estudio?
3. En una ciudad, hay 5 facultades de arquitectura. Si queremos estudiar la opinión de todos los estudiantes de arquitectura sobre un nuevo software de diseño, ¿cuál sería la población?
4. Un estudio quiere analizar la durabilidad de todos los puentes peatonales construidos en una ciudad en los últimos 10 años. ¿Cuál es la población de este estudio?
5. Una empresa de materiales de construcción quiere conocer la opinión de todos sus clientes sobre un nuevo tipo de ladrillo. Si la empresa tiene 500 clientes, ¿cuál es la población en este estudio?

Sesión 1.2 Identifique en los siguientes enunciados los conceptos de Muestra

6. En una ciudad con 100 edificios de oficinas, un arquitecto decide estudiar 15 de ellos para analizar la eficiencia energética. ¿Cuál es la muestra en este estudio?
7. Un ingeniero civil quiere evaluar la resistencia de un tipo de concreto. Aunque hay 500 construcciones que usaron ese concreto, decide tomar muestras de solo 20 construcciones. ¿Cuál es la muestra en este estudio?
8. En una universidad con 10,000 estudiantes, se realiza una encuesta sobre hábitos de estudio y se selecciona a 500 estudiantes al azar para responderla. ¿Cuál es la muestra en este estudio?

9. Una empresa constructora ha edificado 50 edificios residenciales en una región. Para evaluar la satisfacción de los residentes con respecto a la calidad de la construcción, decide hacer una encuesta en 8 de esos edificios. ¿Cuál es la muestra?
10. Un estudio quiere analizar la durabilidad de un tipo de pintura en exteriores. Hay 200 casas en una comunidad que usaron esa pintura, pero el estudio decide evaluar solo 30 de esas casas. ¿Cuál es la muestra en este estudio?

Sesión 1.3 Identifique en los siguientes enunciados los conceptos de Variable

11. En un estudio sobre los tipos de materiales utilizados en la construcción de edificios, se recopilan los siguientes datos: concreto, acero, madera, ladrillo. ¿Qué tipo de variable es esta?
12. Un ingeniero civil está evaluando la resistencia a la compresión de diferentes mezclas de concreto. Registra los siguientes valores en MPa: 25, 30, 28, 32, 29. ¿Qué tipo de variable es esta?
13. En un proyecto arquitectónico, se quiere clasificar los espacios de una casa en: habitaciones, baños, cocina, sala, comedor. ¿Qué tipo de variable es esta clasificación?
14. Un arquitecto está diseñando un edificio y quiere evaluar la cantidad de luz natural que reciben diferentes áreas. Mide la luminosidad en lux y obtiene los siguientes valores: 500 lux, 520 lux, 480 lux, 530 lux. ¿Qué tipo de variable es esta?
15. En un estudio sobre la durabilidad de diferentes tipos de pavimentos en carreteras, se clasifican los pavimentos en: asfalto, concreto, adoquín, grava. ¿Qué tipo de variable es esta clasificación?

Sesión 1.4 Identifique en los siguientes enunciados los conceptos de nivel de medición de los datos

16. Se te presenta una lista de materiales utilizados en una construcción: ladrillo, concreto, madera, acero. ¿Qué tipo de nivel de medición representan estos materiales?

17. En un proyecto de construcción, se registraron los siguientes tipos de suelos: arenoso, arcilloso, limoso, rocoso. ¿Qué tipo de nivel de medición representan estos tipos de suelos?
18. Se te da una lista de estilos arquitectónicos: gótico, moderno, colonial, brutalista. ¿Qué nivel de medición representan estos estilos?
19. En un proyecto de construcción, se clasificaron las zonas según su prioridad: alta, media, baja. ¿Qué nivel de medición representan estas prioridades?
20. Se te da una lista de resistencia de materiales: excelente, buena, regular, pobre. ¿Qué nivel de medición representan estas categorías de resistencia?
21. En un proyecto arquitectónico, se clasificaron las áreas según su importancia: primaria, secundaria, terciaria. ¿Qué nivel de medición representan estas áreas?
22. Se registraron las temperaturas (en grados Celsius) en diferentes zonas de un edificio: 20°C, 22°C, 24°C, 18°C. ¿Qué nivel de medición representan estas temperaturas?
23. Se midieron las inclinaciones (en grados) de diferentes rampas en un proyecto: 5°, 10°, 15°, 20°. ¿Qué nivel de medición representan estas inclinaciones?
24. Se registraron las intensidades lumínicas (en lux) en diferentes áreas de un edificio: 300 lux, 500 lux, 700 lux. ¿Qué nivel de medición representan estas intensidades lumínicas?
25. Se registró la cantidad de concreto (en metros cúbicos) utilizado en diferentes partes de una construcción: 10 m³, 15 m³, 20 m³. ¿Qué nivel de medición representan estas cantidades?
26. Se midió la longitud (en metros) de diferentes vigas en un proyecto: 5 m, 10 m, 15 m. ¿Qué nivel de medición representan estas longitudes?
27. Se registró el tiempo (en días) que tardó la construcción de diferentes partes de un proyecto: 30 días, 60 días, 90 días. ¿Qué nivel de medición representan estos tiempos?

Conclusiones del Capítulo 1

El primer capítulo del libro estableció los fundamentos esenciales para comprender la aplicación de la estadística en la ingeniería civil y la arquitectura. Se introdujeron conceptos clave como población, muestra, parámetro, estadístico y pruebas de hipótesis, cuya correcta comprensión resulta indispensable para la aplicación efectiva de métodos estadísticos en el análisis y diseño de proyectos. Tanto la estadística descriptiva como la inferencial han demostrado ser herramientas fundamentales para examinar datos, hacer inferencias y respaldar decisiones con base en evidencia.

Además, se abordó la clasificación de los datos en cualitativos (nominales y ordinales) y cuantitativos (discretos y continuos), así como sus diferentes niveles de medición (nominal, ordinal, de intervalo y de razón). Se destacó la importancia de reconocer estos tipos y niveles, ya que su correcta identificación permite seleccionar las técnicas de análisis más apropiadas y garantizar una interpretación precisa de los resultados.

El capítulo también presentó el proceso general de análisis estadístico, que incluye la recopilación, organización, limpieza, exploración, análisis descriptivo e inferencial, interpretación y presentación de resultados. Cada una de estas etapas ha sido resaltada como crucial para asegurar que los datos sean manejados de manera eficiente y que los hallazgos obtenidos sean precisos y aplicables en la gestión de proyectos de ingeniería.

Se concluyó que la estadística no solo aporta métodos cuantitativos para evaluar información, sino que también proporciona un marco estructurado para tomar decisiones fundamentadas. La combinación de un sólido entendimiento del análisis estadístico con el conocimiento técnico de la ingeniería civil y la arquitectura permite a los profesionales diseñar estructuras seguras, optimizar recursos y mejorar la eficiencia en la gestión de proyectos.

Capítulo 2

RECOPIACIÓN Y ORGANIZACIÓN DE DATOS

Este capítulo aborda la recopilación y organización de datos como base para el análisis estadístico. Se exploran las diversas fuentes de datos, desde encuestas hasta registros históricos, y la importancia de planificar su recolección para garantizar información relevante y confiable. Además, se presentan métodos de organización como tablas, gráficos y software estadístico para facilitar su interpretación.



La estadística, como disciplina fundamental en la toma de decisiones basadas en datos, comienza con la recopilación y organización de información. Este proceso es esencial para obtener una visión clara y precisa de los fenómenos que se estudian. En este capítulo, exploraremos los métodos y técnicas necesarios para recopilar datos y organizarlos de manera efectiva.

Antes de embarcarse en cualquier análisis estadístico, es necesario determinar de dónde provendrán los datos. Las fuentes de datos pueden variar ampliamente y pueden incluir encuestas, experimentos, registros históricos, observaciones, sensores, fuentes en línea y más. La elección de la fuente de datos dependerá del propósito de la investigación y de la disponibilidad de recursos.

El diseño de la recopilación de datos es un proceso clave. Se debe planificar para garantizar la obtención de información relevante y confiable. Aquí se discuten aspectos como la selección de la muestra, la formulación de preguntas en encuestas, la duración de la recopilación y la logística de campo.

Una vez que los datos han sido recopilados, es preciso organizarlos de manera que sean comprensibles y útiles. En esta sección, se explican métodos de organización de datos, como tablas, gráficos, hojas de cálculo y software estadístico.

2.1. Concepto de recolección de datos

La recolección de datos es el uso de técnicas e instrumentos para recopilar información acerca de un tema específico que se está investigando. Este proceso es importante para asegurar que los resultados obtenidos sean confiables y útiles. Según Mazhar (2021), los datos pueden clasificarse como primarios, aquellos recopilados directamente de fuentes originales como encuestas y experimentos, o secundarios, obtenidos de fuentes previamente publicadas como libros o revistas.

Además, esta etapa desempeña un papel esencial en el análisis de sistemas de información, ya que de ella depende la calidad del producto que se desea desarrollar. Schadewitz y Blevins (1997) señalan que la selección y recopilación adecuada de datos facilita

un análisis posterior más efectivo y alineado con los objetivos de la investigación.

2.2. Técnica de recolección de datos

La recopilación de datos es un paso fundamental en cualquier investigación estadística, ya que proporciona la materia prima necesaria para llevar a cabo análisis significativos. La elección de las técnicas de recolección de datos adecuadas es fundamental para obtener información precisa y confiable. Aquí, exploramos algunas de las técnicas más comunes utilizadas en estadística:

2.2.1. Encuestas

Las encuestas son un método ampliamente utilizado para recopilar datos de una muestra representativa de una población. Pueden ser realizadas de manera presencial, telefónica, por correo o en línea. Las preguntas pueden ser cerradas (con opciones predefinidas) o abiertas (donde los encuestados pueden responder en sus propias palabras). Se obtiene información sobre las necesidades y preferencias de usuarios o clientes. Difiere de un censo en donde toda la población es estudiada.

2.2.2. Experimentos

Los experimentos implican la manipulación deliberada de una o más variables independientes para observar su efecto en una variable dependiente. Los datos se recopilan bajo condiciones controladas y se utilizan para establecer relaciones causales.

2.2.3. Observación Directa

Esta técnica implica la observación y registro de eventos o comportamientos tal como ocurren en un entorno natural. Puede ser útil en estudios de campo, como la observación de patrones de tráfico o el comportamiento de los consumidores en una tienda. Permite al analista determinar que se está haciendo, como se está haciendo, quien lo hace, cuando se lleva a cabo, cuanto tiempo toma, dónde se hace y por qué se hace.

2.2.3.1. Preparación para la Observación

Determinar y definir las variables que va a observarse.

Estimar el tiempo necesario de observación.

Explicar a las personas que van a ser observadas lo que se va a hacer y las razones para ello.

2.2.4. Registros y Bases de Datos

En muchas investigaciones, los datos ya existen en forma de registros, archivos o bases de datos. Estos registros pueden incluir datos demográficos, financieros, médicos, etc. Los investigadores pueden acceder a estos datos para su análisis.

2.2.5. Entrevistas

Las entrevistas implican la comunicación directa entre el investigador y el entrevistado. Pueden ser estructuradas (con preguntas predeterminadas) o no estructuradas (conversaciones abiertas). Las entrevistas son útiles cuando se requiere una comprensión más profunda de las respuestas.

2.2.6. Cuestionarios en Línea

Con la creciente prevalencia de la tecnología, los cuestionarios en línea se han convertido en una forma efectiva de recopilar datos. Plataformas en línea permiten a los encuestados responder preguntas de manera eficiente y a los investigadores acceder a los resultados de manera conveniente.

2.2.7. Análisis de Contenido

Esta técnica se utiliza para analizar el contenido de documentos escritos, medios audiovisuales o datos en línea. Los investigadores buscan patrones y temas dentro de los datos para extraer información relevante.

2.2.8. Sensores y Dispositivos Inteligentes

Con avances en la tecnología, se pueden utilizar sensores y dispositivos inteligentes para recopilar datos en tiempo real. Esto se

aplica en campos como la monitorización de la salud, el seguimiento de datos medioambientales y más.

2.2.9. Grupos de Discusión

Los grupos de discusión involucran a un pequeño grupo de participantes que discuten un tema en profundidad. Esto es útil para obtener perspectivas cualitativas sobre un tema.

La elección de la técnica de recolección de datos adecuada depende del objetivo de la investigación, el alcance del estudio y los recursos disponibles. Es importante diseñar el proceso de recopilación de datos de manera que minimice el sesgo y garantice la fiabilidad y la validez de los resultados. Una vez que se recopilan los datos, el siguiente paso es su organización y análisis, lo que permitirá obtener conclusiones significativas.

2.3. Instrumentos de Recolección de Datos en Estadística

La elección de los instrumentos de recolección de datos es fundamental en estadística, ya que determina la calidad y la confiabilidad de la información que se recopila. Los instrumentos son las herramientas o métodos utilizados para obtener datos de una muestra o población. A continuación se describen algunos instrumentos comunes utilizados en la recopilación de datos en estadística:

2.3.1. Cuestionarios y Encuestas

Los cuestionarios y encuestas son instrumentos ampliamente utilizados en la recopilación de datos cuantitativos. Pueden ser en papel o digitales y contienen una serie de preguntas cerradas o abiertas que los participantes completan. Los cuestionarios estructurados permiten la recopilación de datos estandarizados, lo que facilita el análisis posterior.

2.3.2. Entrevistas

Las entrevistas implican una interacción directa entre el investigador y el entrevistado. Pueden ser estructuradas, con un conjunto de

preguntas predefinidas, o no estructuradas, permitiendo una conversación abierta. Las entrevistas son útiles para recopilar datos cualitativos y obtener una comprensión más profunda de las respuestas.

2.3.3. Observación Directa

La observación implica la recolección de datos a través de la observación de eventos o comportamientos en un entorno natural. Los investigadores observan y registran lo que sucede sin intervención directa. Esta técnica es común en estudios de campo y es útil para recopilar datos cualitativos.

2.3.4. Registro de Datos

En algunos casos, los datos ya están registrados en forma de registros, bases de datos o archivos. Los investigadores pueden acceder a estos registros para su análisis. Esto es común en la investigación médica, financiera y de registros administrativos.

2.3.5. Escalas de Medición

Las escalas de medición son instrumentos diseñados para medir características específicas de los objetos de estudio. Pueden incluir escalas de opinión, escalas de actitud, escalas de dolor y muchas otras. Estas escalas asignan valores numéricos a las características y permiten la cuantificación de datos.

2.3.6. Sensores y Dispositivos Inteligentes

En la era de la tecnología, se utilizan sensores y dispositivos inteligentes para recopilar datos en tiempo real. Estos dispositivos pueden medir variables como temperatura, presión, movimiento, ubicación GPS y más. Son comunes en campos como la monitorización de la salud, la ciencia ambiental y la ingeniería.

2.3.7. Grupos de Discusión

Los grupos de discusión son una técnica cualitativa que involucra a un pequeño grupo de participantes en una discusión abierta sobre

un tema específico. El investigador facilita la conversación y registra las opiniones y perspectivas de los participantes.

2.3.8. Análisis de Contenido

El análisis de contenido es un enfoque cualitativo que implica el estudio sistemático de documentos escritos, medios audiovisuales y datos en línea. Los investigadores identifican patrones, temas y categorías en el contenido.

La elección del instrumento de recolección de datos depende del objetivo de la investigación, el tipo de datos necesarios y los recursos disponibles.

Es importante seleccionar la técnica que mejor se adapte a la investigación en cuestión y diseñar el proceso de recopilación de datos de manera que minimice el sesgo y garantice la confiabilidad y la validez de los resultados.

Los instrumentos permiten registrar la información recolectada y facilitan la toma de decisiones y cálculos estadísticos. Se puede hacer de manera presencial o no presencial.

2.4. Organización de los datos

La organización de los datos en estadística es un paso esencial para facilitar el análisis y la interpretación de la información recopilada. A continuación se describen algunas de las principales técnicas y métodos de organización de datos.

2.4.1. Tablas de Frecuencias

Una tabla de frecuencia es una forma común de organizar datos cuantitativos y categóricos. Las tablas de frecuencia proporcionan una visión general de la distribución de los datos.

Son herramientas estadísticas utilizadas para organizar y resumir datos cuantitativos o cualitativos, mostrando la cantidad de veces que ocurre cada valor o rango de valores en un conjunto de datos. Estas tablas permiten visualizar la distribución de los datos de manera clara y concisa.

Ejemplos:

En la Tabla 1 se muestra los datos de estilos arquitectónicos en un conjunto de edificios históricos, datos categóricos.

En la Tabla 2 se visualiza que estás evaluando la resistencia a la compresión de muestras de concreto utilizadas en un proyecto de construcción.

Tabla 1

Datos de Edificios por Estilo Arquitectónico

Estilo Arquitectónico	Frecuencia Absoluta	Frecuencia Relativa (%)	Frecuencia Acumulada	Frecuencia Relativa Acumulada (%)
Neoclásico	4	40.0	4	40.0
Moderno	4	40.0	8	80.0
Gótico	2	20.0	10	100.0
Barroco	2	20.0	10	100.0

Tabla 2

Datos de Muestras de Concreto Resistencia en Mpa

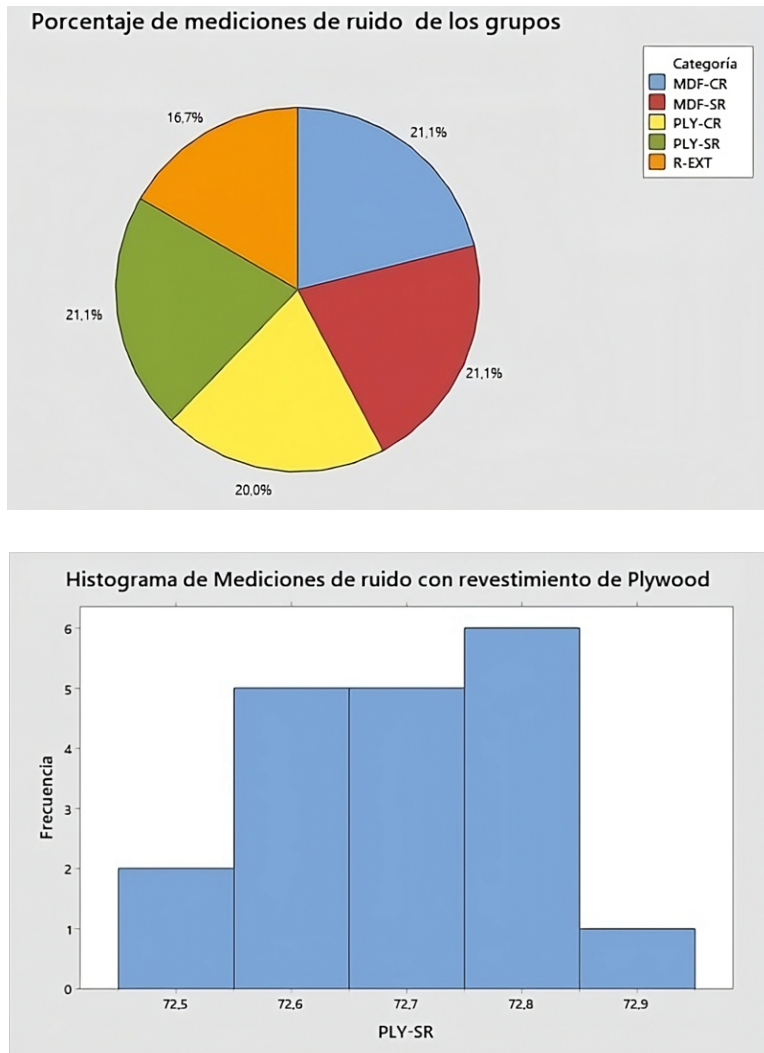
Muestra	Resistencia (MPa)
1	25
2	30
3	28
4	30
5	35
6	25
7	28
8	30
9	35
10	30

2.4.2. Gráficos y Diagramas

Los gráficos y diagramas son herramientas visuales para representar datos. Los gráficos de barras, los histogramas y los gráficos circulares son útiles para visualizar la distribución de datos categóricos o cuantitativos.

Los gráficos de dispersión son ideales para mostrar la relación entre dos variables cuantitativas. En la Figura 1 se muestran los diferentes tipos de gráficos y diagramas.

Figura 1
Tipos de gráficos



2.4.3. Matrices de Datos

Las matrices de datos organizan los datos en filas y columnas, con cada fila representando una observación y cada columna una variable.

Este formato es especialmente útil para el análisis de datos multivariados.

En la Tabla 3 se encuentran un ejemplo de las matrices.

Tabla 3

Matrices de datos ruidos de revestimiento con Playwood y paja toquilla

PLY-SR	PLY-CR	MDF-SR	MDF-CR	R-EXT	RUIDO	GRUPO
72,9	69,9	74,2	71,2	92,5	72,9	PLY-SR
72,8	69,8	74,4	71,2	92,9	72,8	PLY-SR
72,6	69,6	74,4	71,4	94,2	72,6	PLY-SR
72,7	70,1	74,3	71,3	94,4	72,7	PLY-SR
72,5	70,1	73,7	71,5	94,5	72,5	PLY-SR
72,8	69,9	74,1	71,3	94,2	72,8	PLY-SR
72,8	69,9	73,9	71,6	94,3	72,8	PLY-SR
72,8	70,1	74,2	70,5	94,1	72,8	PLY-SR
72,7	70,3	74,3	71,3	94,1	72,7	PLY-SR

2.4.4. Clasificación De Tablas de Frecuencias

Las tablas de frecuencias varían según la forma en que los datos se agrupan y presenten. En función de la agrupación de datos, los tipos de tablas de frecuencias más comunes son:

2.4.4.1. Tabla de Frecuencias para Datos No Agrupados o Discretos simples

Se utiliza cuando se tienen datos individuales o valores discretos, como la cantidad de estudiantes en una clase o la edad exacta de personas. En esta tabla, cada valor se lista individualmente, junto con su frecuencia absoluta y, opcionalmente, la frecuencia relativa.

Ejemplo:

Supongamos que estamos analizando la cantidad de puertas en diferentes tipos de edificios en una muestra de 20 edificios.

Cantidad de puertas por edificio: 2, 3, 4, 3, 5, 2, 4, 3, 5, 6, 2, 3, 4, 5, 6, 2, 3, 4, 5, 4.

En la Tabla 4 se muestran los datos individuales y las frecuencias.

Tabla 4

Tabla de frecuencia de la cantidad de puertas

Cantidad de Puertas	Frecuencia Absoluta	Frecuencia Relativa (%)	Frecuencia Acumulada	Frecuencia Relativa Acumulada (%)
2	4	20	4	20
3	5	25	9	45
4	5	25	14	70
5	4	20	18	90
6	2	10	20	100.0

2.4.4.2. Tabla de Frecuencias para Datos Agrupados en Intervalos o Continuos

Se utiliza cuando se tienen datos continuos y es necesario agruparlos en intervalos (clases) para facilitar la presentación y el análisis. En esta tabla, los datos se agrupan en intervalos, y se muestra la frecuencia absoluta (n) y, opcionalmente, la frecuencia relativa (f) para cada intervalo.

Ejemplo:

Supongamos que estamos analizando la altura (en metros) de 30 edificios en una ciudad. Los datos de altura podrían verse así:

Alturas de edificios (en metros):

22.5, 24.3, 28.7, 32.1, 25.6, 27.8, 30.2, 33.4, 29.1, 31.7, 23.4, 24.9, 26.7, 28.5, 30.4, 32.8, 34.1, 35.2, 25.3, 27.6, 29.4, 31.2, 33.5, 34.9, 26.3, 28.2, 29.7, 31.8, 33.0, 35.6.

En la Tabla 5 se muestra en la primera columna los intervalos que se agrupan los 30 datos.

Tabla 5

Tabla de frecuencias de las alturas de edificios

Intervalo de Altura (m)	Frecuencia Absoluta	Frecuencia Relativa (%)	Frecuencia Acumulada	Frecuencia Relativa Acumulada (%)
22.5 - 24.9	4	13	4	13
25.0 - 27.4	5	17	9	30
27.5 - 29.9	7	23	16	53
30.0 - 32.4	6	20	22	73
32.5 - 34.9	5	17	27	90
35.0 - 37.4	3	10	30	100

2.4.4.3. Tabla de Frecuencias para Datos Categóricos

Se emplea para datos categóricos o cualitativos, como el género, la preferencia de productos, etc. Cada categoría se presenta junto con su frecuencia absoluta y, opcionalmente, la frecuencia relativa.

Ejemplo:

Supongamos que estamos analizando los estilos arquitectónicos de 30 edificios en una ciudad. Los estilos arquitectónicos pueden ser

Clásico, Moderno, Colonial, y Art Decó.

Estilos arquitectónicos de los edificios:

Clásico, Moderno, Colonial, Moderno, Art Decó, Clásico, Colonial, Moderno, Art Decó, Clásico, Colonial, Moderno, Clásico, Colonial, Art Decó, Clásico, Moderno, Colonial, Art Decó, Moderno, Clásico, Moderno, Art Decó, Clásico, Colonial, Moderno, Clásico, Colonial, Art Decó, Moderno.

La Tabla 6 se muestra que el estilo Moderno es el más prevalente entre los edificios analizados, representando el 30% del total de la muestra. Le sigue el estilo Clásico, que constituye el 27% de los edificios, lo que indica que estos dos estilos son los más dominantes en la ciudad. El estilo Colonial es también significativo, con un 23% de los edificios, mientras que el estilo Art Decó es el menos común, aunque aún representa una parte considerable del paisaje arquitectónico con un 20% del total. Esta distribución sugiere una diversidad moderada en los estilos arquitectónicos de la ciudad, con una ligera inclinación hacia lo moderno y lo clásico. Estos resultados podrían reflejar tendencias en las preferencias arquitectónicas actuales y en la evolución histórica del entorno construido.

Tabla 6

Tabla de frecuencias de estilos arquitectónicos

Estilo Arquitectónico	Frecuencia Absoluta	Frecuencia Relativa (%)	Frecuencia Acumulada	Frecuencia Relativa Acumulada (%)
Clásico	8	27	8	27
Moderno	9	30	17	57
Colonial	7	23	24	80
Art Decó	6	20	30	100.0

2.4.4.4. Tabla de Frecuencias Cruzadas o de Contingencia

Se utiliza para resumir dos variables categóricas. En esta tabla, se muestra la frecuencia absoluta de las combinaciones de categorías de ambas variables.

La elección del tipo de tabla de frecuencias dependerá de la naturaleza de los datos y de los objetivos del análisis. La clave es

seleccionar el formato que mejor se adapte a tus datos y que proporcione una representación efectiva de la información que estás estudiando.

La tabla de frecuencias es una herramienta fundamental en estadística que se utiliza para organizar y resumir datos, especialmente cuando se trata de datos cuantitativos. Su objetivo es proporcionar una visión general de la distribución de datos y resaltar patrones y características clave. Aquí se describen los componentes y el proceso para crear una tabla de frecuencias.

Ejemplo:

Supongamos que estamos analizando la relación entre el estilo arquitectónico y el material de construcción de 30 edificios en una ciudad. Los estilos arquitectónicos pueden ser Clásico, Moderno, Colonial, y Art Decó. Los materiales de construcción pueden ser Concreto, Madera, y Ladrillo.

Cada celda de la Tabla 7 muestra cuántos edificios corresponden a una combinación específica de estilo arquitectónico y material de construcción.

Tabla 7

Tabla de frecuencias estilo arquitectónico vs material de construcción

Estilo Arquitectónico	Concreto	Madera	Ladrillo	Total
Clásico	3	2	3	8
Moderno	6	1	2	9
Colonial	1	4	2	7
Art Decó	2	3	1	6
Total	12	10	8	30

La tabla de contingencia sugiere varias relaciones interesantes entre el estilo arquitectónico y el material de construcción de los edificios. El estilo Moderno es el que más frecuentemente utiliza concreto como material principal, mientras que el estilo Colonial muestra una preferencia por la madera. El estilo Clásico es relativamente equilibrado en su uso de concreto, madera y ladrillo, mientras que Art Decó tiene una ligera preferencia por la madera.

Estos patrones pueden indicar tendencias en las elecciones de materiales según el estilo arquitectónico en esta muestra de edificios. La tabla de contingencia es una herramienta poderosa para identificar posibles asociaciones entre variables categóricas, lo que puede ser útil para tomar decisiones informadas en diseño, construcción y planificación urbana.

2.5. Componentes de una Tabla de Frecuencias

Una tabla de frecuencias es una herramienta estadística fundamental que organiza y resume un conjunto de datos, mostrando cuántas veces ocurre cada valor o grupo de valores en el conjunto de datos. Es especialmente útil para interpretar y visualizar la distribución de los datos, facilitando el análisis de patrones y tendencias. Dependiendo del tipo de datos, las tablas de frecuencias pueden ser simples, agrupadas, categóricas, o cruzadas, cada una adaptada a la naturaleza de la información que se desea analizar. Además, su correcta estructuración permite una mejor interpretación de la variabilidad y concentración de los datos.

Los principales componentes de una tabla de frecuencias son:

2.5.1. Clases o Intervalos

En una tabla de frecuencias, los datos cuantitativos se agrupan en clases o intervalos. Cada intervalo abarca un rango específico de valores. La elección de los intervalos es importante y debe ser representativa de la distribución de los datos.

2.5.2. Frecuencia Absoluta (f_a)

La frecuencia absoluta en una tabla de frecuencias representa la cantidad de observaciones que caen dentro de cada intervalo. Es simplemente el recuento de datos en ese intervalo.

2.5.3. Frecuencia Relativa (f_r).

La frecuencia relativa se calcula dividiendo la frecuencia absoluta en un intervalo entre el número total de observaciones (n). Es decir, $f = n/\text{número total de observaciones}$.

2.5.4. Frecuencia Acumulada (Fa)

La frecuencia acumulada es la suma de las frecuencias absolutas acumuladas desde el primer intervalo hasta el intervalo actual. Es útil para analizar la distribución acumulada de datos.

2.5.5. Frecuencia Relativa Acumulada (Fr)

La frecuencia relativa acumulada se calcula dividiendo la frecuencia acumulada en un intervalo entre el número total de observaciones. Es decir, $F = N/\text{Número total de observaciones}$.

2.6. Creación de tablas de frecuencias para datos simples.

Crear una tabla de frecuencias para datos simples implica organizar y resumir la frecuencia de cada valor individual en una variable discreta.

A continuación, te proporciono el proceso paso a paso para realizar una tabla de frecuencias de datos simples:

Paso 1: Recopilación de Datos

Reúne todos los datos que deseas analizar. Asegúrate de que estos datos sean discretos o categóricos, lo que significa que cada valor sea único y no existan intervalos.

Paso 2: Ordena los Datos

Si los datos no están ya ordenados, organízalos en orden ascendente o descendente. Esto facilitará la identificación de duplicados y la asignación de frecuencias.

Paso 3: Identificación de Valores Únicos

Para cada valor único en tus datos, cuenta cuántas veces aparece en tu conjunto de datos. Esto se conoce como la *frecuencia absoluta* de cada valor.

Paso 4: Creación de la Tabla de Frecuencias

Diseña una tabla con al menos tres columnas:

La primera columna debe contener los valores únicos de tus datos.

La segunda columna debe mostrar la frecuencia absoluta de cada valor.

Opcionalmente, puedes agregar una tercera columna para la frecuencia relativa, que se calcula dividiendo la frecuencia absoluta de cada valor entre el total de observaciones. La frecuencia relativa se expresa como un porcentaje.

Paso 5: Calcular la Frecuencia Acumulada

Si deseas, puedes agregar una columna adicional para la frecuencia acumulada. Esto implica sumar la frecuencia absoluta de cada valor actual a la frecuencia acumulada de los valores anteriores. La frecuencia acumulada proporciona una visión acumulativa de la distribución de datos.

Una vez completada la tabla de frecuencias, utilízala para comprender la distribución de tus datos. Puedes observar los valores más comunes, identificar valores atípicos y obtener una visión general de la variabilidad de tus datos.

La tabla de frecuencias es una herramienta esencial en estadística para resumir datos discretos o categóricos. Proporciona una representación clara de la frecuencia de cada valor en los datos, lo que es fundamental para el análisis y la toma de decisiones.

2.7. Ejercicio de construcción de tablas de frecuencias para datos simples

Un ingeniero civil registra la cantidad de proyectos completados por sus empleados en un año: 5, 6, 5, 7, 6, 5, 7, 6.

Con estos datos, elabora una tabla de frecuencias simples.

Paso 1: Organizar los Datos

Primero, organiza los datos en orden ascendente para facilitar el análisis. Datos organizados: 5, 5, 5, 6, 6, 6, 7, 7

Paso 2: Identificar los Valores Únicos

Identifica los valores únicos o distintos en el conjunto de datos. En este caso, los valores son 5, 6 y 7.

Paso 3: Contar las Frecuencias (fa)

Cuenta cuántas veces aparece cada valor en el conjunto de datos. Esta será la frecuencia absoluta (f).

- Valor 5: Aparece 3 veces
- Valor 6: Aparece 3 veces
- Valor 7: Aparece 2 veces

Paso 4: Calcular la Frecuencia Relativa (fr)

La frecuencia relativa (f/n) es la proporción de la frecuencia absoluta de cada valor respecto al total de datos. El total de datos (n) es 8.

Frecuencia Relativa del valor 5: $\frac{3}{8} = 0.375$

Frecuencia Relativa del valor 6: $\frac{3}{8} = 0.375$

Frecuencia Relativa del valor 7: $\frac{2}{8} = 0.25$

Paso 5: Calcular la Frecuencia Acumulada (Fa)

La frecuencia acumulada (Fa) es la suma de las frecuencias absolutas hasta el valor correspondiente.

Frecuencia acumulada para el valor 5:
3 (solo el valor 5)

Frecuencia acumulada para el valor 6:
3 (del valor 5) + 3 (del valor 6) = 6

Frecuencia acumulada para el valor 7:
6 (de los valores anteriores) + 2 (del valor 7) = 8

Paso 6: Calcular la Frecuencia relativa Acumulada (Fr)

La frecuencia acumulada relativa (Fr) es la suma de las frecuencias relativas hasta el valor correspondiente.

Frecuencia acumulada para el valor 5:
 0.375 (solo el valor 5)

Frecuencia acumulada para el valor 6:
 0.75 (del valor 5) + 0.375 (del valor 6) = 0.75

Frecuencia acumulada para el valor 7:
 0.75 (de los valores anteriores) + 0.25 (del valor 7) = 1

La Tabla 8 se muestra cuántos proyectos completaron los empleados en un año, así como las frecuencias relativas y acumuladas. Por ejemplo, 3 empleados completaron 5 proyectos, lo que representa el 37.5% de los empleados.

Tabla 8

Tabla de frecuencias de proyectos completados

Proyectos completados	Frecuencia Absoluta	Frecuencia Relativa	Frecuencia Acumulada	Frecuencia Relativa Acumulada
5	3	0.375 (37.5%)	3	0.375 (37.5%)
6	3	0.375 (37.5%)	6	0.75 (75%)
7	2	0.25 (25%)	8	1 (100%)
Total	6	1.0 (100%)		

2.8. Creación de tablas de frecuencias para datos agrupados

Crear una tabla de frecuencias para datos agrupados implica organizar y resumir la frecuencia de observaciones en intervalos o clases en lugar de valores individuales.

Aquí está el proceso paso a paso para hacer una tabla de frecuencias para datos agrupados:

Paso 1: Recopilación de Datos

Reúne los datos que deseas analizar. Asegúrate de que los datos sean cuantitativos y continuos, lo que significa que puedan agruparse en intervalos.

Paso 2: Definición de Intervalos (Clases)

Decide cómo agrupar los datos en intervalos o clases. Los intervalos deben ser mutuamente excluyentes y representativos de la distribución de los datos. La anchura de los intervalos es un factor importante y debe elegirse cuidadosamente. Los intervalos pueden superponerse, pero no deben dejar brechas entre ellos.

Para decidir el número de intervalos en una tabla de frecuencias, existen varias fórmulas y métodos comunes. Aquí te presento algunos de los más utilizados:

- **Regla de Sturges:**

Fórmula:

$$k = 1 + \log_2(n), \text{ se puede escribir como } k = 1 + 3.3 \log (n)$$

Donde:

k es el número de intervalos

n es el número total de datos.

Esta fórmula sugiere un número adecuado de intervalos basado en el tamaño de la muestra.

- **Regla de Rice:**

Fórmula:

$$k = \sqrt{n}$$

Similar a la regla de Sturges, pero se basa en la raíz cuadrada del número total de datos. Es útil cuando se tienen grandes conjuntos de datos.

- **Método de Ajuste Manual**

Basado en la interpretación de los datos y el contexto. A veces, es necesario ajustar el número de intervalos manualmente para que la tabla sea más útil y clara.

Paso 3: Determinación del Rango de Datos

Encuentra el rango de los datos, que es la diferencia entre el valor más grande y el valor más pequeño en tus datos.

$$R = \text{Máximo} - \text{Mínimo}$$

Paso 4: Cálculo del Ancho del Intervalo (Amplitud)

Decide el ancho de cada intervalo. Una regla general es utilizar entre 5 y 20 intervalos, según la cantidad de datos y la distribución.

Puedes calcular la amplitud (ancho) dividiendo el rango de datos por el número de intervalos deseado.

$$\text{Amplitud} = R/k$$

Donde:

R es Rango de Datos (calculado en el Paso 3)

k es Número de Intervalos (calculado en el Paso 2).

Paso 5: Creación de la Tabla de Frecuencias

Diseña una tabla con al menos cinco columnas:

- La primera columna muestra los límites inferiores de los intervalos.
- La segunda columna muestra los límites superiores de los intervalos.
- La tercera columna muestra el punto medio de cada intervalo (opcional).
- La cuarta columna muestra la frecuencia absoluta, que es el número de observaciones que caen en cada intervalo.
- La quinta columna puede mostrar la frecuencia relativa, que se calcula dividiendo la frecuencia absoluta entre el número total de observaciones.

Paso 6: Contabilización de Observaciones en Intervalos

Para cada intervalo, cuenta cuántas observaciones de datos caen dentro de ese rango. Esto se convierte en la frecuencia absoluta para ese intervalo.

Paso 7: Cálculo de Frecuencias Relativas

Si decides incluir la frecuencia relativa en tu tabla, divide la frecuencia absoluta en cada intervalo entre el número total de observaciones. Esto proporcionará la frecuencia relativa para cada intervalo.

Paso 8: Presentación y Análisis de los Resultados

Una vez que se completa la tabla de frecuencias, utilízala para comprender la distribución de tus datos agrupados. Esto te permitirá identificar patrones, tendencias y características clave de los datos y facilitará el análisis y la interpretación de los resultados.

2.9. Ejercicio de construcción de tablas de frecuencias para datos agrupados

A continuación, te presento ejercicios relacionados con cómo armar tablas de distribución de frecuencia desde datos iniciales, incluyendo el cálculo del rango, número de clases, ancho de clase, y la tabla con frecuencia absoluta, relativa, absoluta acumulada y relativa acumulada.

Ejercicio 1:

Un arquitecto ha registrado las áreas (en m²) de 10 salas de conferencias en un centro de convenciones: 100, 105, 102, 110, 108, 105, 103, 107, 106, 104. Elabora una tabla de distribución de frecuencia para estos datos.

Paso 1: Recopilación de Datos

Reúne los datos y ordénalos si es necesario:

Datos: 100, 105, 102, 110, 108, 105, 103, 107, 106, 104

Paso 2: Determina el Rango de Datos

Valor mínimo = 100

Valor máximo = 110

Rango = Valor máximo - Valor mínimo.

Rango: $110 - 100 = 10$.

Paso 3: Decide el Número de Intervalos

Usa una fórmula o método para decidir el número de intervalos. Para este ejemplo, utilizaremos la Regla de Sturges:

Número de datos (n) = 10

Número de clases (usando la regla de Sturges): $k=1+3.32 \times \log(10) \approx 4.32$

Redondeamos el número de intervalos a 4 para simplificar.

Paso 4: Calcula el Tamaño del Intervalo

Tamaño del intervalo = Rango / Número de intervalos

Ancho de clase: $10 / 4 = 2.5$

Redondeamos el tamaño del intervalo a 3 para facilitar la construcción de la tabla.

Paso 5: Define los Intervalos

Empieza desde el valor mínimo y usa el tamaño del intervalo para definir los límites:

Intervalo 1: 100 - 103

Intervalo 2: 103 - 106

Intervalo 3: 106 - 109

Intervalo 4: 109 - 112

Paso 6: Cuenta las Frecuencias

Cuenta cuántos datos caen en cada intervalo:

Intervalo 100 - 103: 100, 102 (2 datos)

Intervalo 103 - 106: 105, 105, 103, 104 (4 datos)

Intervalo 106 - 109: 106, 107, 108 (3 datos)

Intervalo 109 - 112: 110 (1 dato)

Paso 7: Elabora la Tabla de Frecuencias

Esta Tabla 9 se muestra la distribución de las áreas de las salas de conferencias en los intervalos definidos.

Tabla 9

Tabla de frecuencias del área

Área (m2)	Frecuencia Absoluta	Frecuencia Relativa	Frecuencia Acumulada	Frecuencia Relativa Acumulada
[100-103)	2	0,2	2	0,2
[103-106)	4	0,4	6	0,6
[106-109)	3	0,30	9	0,9
[109-112)	1	0,1	10	1
Total	10	1.0 (100%)		

Ejercicios del Capítulo 2 a desarrollar

1. En un estudio sobre los tipos de vehículos que pasan por un puente durante una hora, se registran los siguientes datos de 12 vehículos: coche, camión, coche, bicicleta, coche, camión, coche, coche, moto, camión, bicicleta, coche. Elabora una tabla de frecuencias para datos simples.
2. Un arquitecto realiza un conteo de las ventanas de diferentes tamaños en un edificio de 15 ventanas: grande, mediana, grande, pequeña, mediana, grande, pequeña, grande, mediana, grande, pequeña, grande, pequeña, mediana, grande. Representa estos datos en una tabla de frecuencias con las frecuencias solicitadas.
3. Un arquitecto realiza un estudio sobre los tipos de materiales utilizados en 50 construcciones recientes en la ciudad. Los resultados son los siguientes: concreto, ladrillo, concreto, madera, concreto, acero, ladrillo, concreto, madera, acero, entre otros. Se elaboró la tabla de distribución de frecuencias mostrada a continuación. Determina qué material es el más utilizado.
4. Un ingeniero civil ha registrado la resistencia (en MPa) de 8 muestras de concreto: 30, 32, 31, 30, 33, 32, 31, 34. Elabora una tabla de distribución de frecuencia con intervalos para estos datos.
5. Un arquitecto ha registrado la altura (en metros) de 6 edificios: 20, 25, 23, 20, 27, 25. Elabora una tabla de distribución de frecuencia con intervalos para estos datos.
6. Un arquitecto está estudiando la altura (en metros) de diferentes edificios en una ciudad. Los datos son: 15, 20, 18, 15, 22, 20, 19, 21, entre otros. Se elaboró la Tabla 10. de distribución de frecuencias agrupadas mostrada a continuación. Determina cuántos edificios tienen una altura superior a 18 metros.
7. Un arquitecto está investigando el área (en m^2) de diferentes salones de clases en una universidad. Los datos son: 60, 65, 62, 60, 68, 65, 63, 67, y más, representados en la Tabla 11 de frecuencia. Determina cuántos salones tienen un área entre $62 m^2$ y $67 m^2$.

Tabla 10*Tabla de frecuencias altura de los edificios de la ciudad.*

Altura (m)	Frecuencia Absoluta	Frecuencia Relativa	Frecuencia Acumulada	Frecuencia Relativa Acumulada
15-16	10	0.25	10	0.25
17-18	8	0.20	18	0.45
19-20	12	0.30	30	0.75
21-22	10	0.25	40	1.00
Total	40	1.0 (100%)		

Tabla 11*Tabla de frecuencias áreas de los salones de clases*

Área (m2)	Frecuencia Absoluta	Frecuencia Relativa	Frecuencia Acumulada	Frecuencia Relativa Acumulada
60-61	10	0.25	10	0.25
62-63	8	0.20	18	0.45
64-65	12	0.30	30	0.75
66-67	10	0.25	40	1.00
Total	40	1.0 (100%)		

Conclusiones del Capítulo 2

El capítulo dos permitió comprender la importancia fundamental de la recopilación y organización de datos en la ingeniería civil y la arquitectura, estableciendo las bases para un análisis estadístico sólido. Estos procesos no solo garantizan la precisión y fiabilidad de la información, sino que también facilitan la toma de decisiones fundamentadas en evidencia, optimizando el desarrollo de proyectos en ambas disciplinas.

Se evidenció que la selección de la técnica de recolección de datos adecuada es un factor determinante para obtener información relevante. En ingeniería civil, encuestas, experimentos y observaciones directas han demostrado ser herramientas esenciales para evaluar la calidad de las construcciones y la resistencia de los materiales. En arquitectura, la recopilación de datos a través de entrevistas y sensores inteligentes ha permitido diseñar espacios más eficientes y adaptados a las necesidades de los usuarios. La combinación de datos cuantitativos y cualitativos, obtenidos con instrumentos de medición precisos, ha enriquecido la evaluación de aspectos estructurales y estéticos en los proyectos.

La organización de los datos ha sido otro aspecto clave abordado en el capítulo. Se resaltó que el uso de tablas, gráficos y diagramas no solo facilita la interpretación de la información, sino que también permite detectar patrones, tendencias y relaciones dentro de los datos. Herramientas como histogramas y diagramas de dispersión han resultado fundamentales para la visualización efectiva de la información, lo que mejora la comunicación de hallazgos a todas las partes involucradas en un proyecto.

Además, se destacó la importancia de estructurar los datos de manera clara y ordenada para garantizar su utilidad en el análisis estadístico posterior. La correcta organización de la información ha permitido a los profesionales reducir errores, mejorar la eficiencia en el manejo de recursos y fortalecer la toma de decisiones en el diseño y construcción de infraestructuras.

En definitiva, este capítulo reafirmó que una recolección y organización de datos bien ejecutada es crucial para la ingeniería civil y la arquitectura. La correcta aplicación de estos principios ha demostrado ser un factor clave en la creación de proyectos más seguros, funcionales y eficientes, alineados con las normativas vigentes y las expectativas de los usuarios. La capacidad de gestionar datos de manera efectiva representa una ventaja competitiva en un entorno donde la precisión y la optimización de recursos son determinantes para el éxito de cualquier iniciativa.

Capítulo 3

MEDIDAS DE TENDENCIA CENTRAL

Este capítulo aborda el uso de las estadísticas descriptivas para analizar y resumir datos en ingeniería, facilitando la toma de decisiones. Se explican medidas de centralidad como la media, mediana y moda, junto con su importancia en la interpretación de datos. Además, se presentan herramientas gráficas como histogramas y boxplots para visualizar distribuciones y relaciones entre variables.



3.1. Introducción

En este capítulo, exploraremos el poder y la utilidad de las estadísticas descriptivas en el contexto de la interpretación y análisis de datos en diversas áreas de la ingeniería. Las estadísticas descriptivas son una herramienta fundamental que nos permite resumir, organizar y presentar de manera significativa grandes conjuntos de datos, proporcionando insights valiosos para la toma de decisiones informadas.

Examinaremos los conceptos básicos de las estadísticas descriptivas, incluyendo medidas de centralidad como la media, la mediana y la moda. Estas medidas nos permiten comprender la tendencia central de nuestros datos dentro de ellos, aspectos necesarios para entender la naturaleza de nuestros fenómenos estudiados.

Exploraremos herramientas gráficas como histogramas, diagramas de caja (boxplots) y gráficos de dispersión, que complementan las medidas descriptivas numéricas al proporcionar una representación visual de la distribución y la relación entre variables en nuestros datos.

3.2. Media Aritmética

La media aritmética es una medida estadística que representa el valor típico de un conjunto de datos numéricos. Se calcula sumando todos los valores de los datos y dividiendo esta suma por el número total de elementos en el conjunto. Es una medida de tendencia central que proporciona información sobre el valor medio de un conjunto de datos.

Además, la media aritmética es una herramienta sencilla pero esencial en estadística, particularmente útil cuando los datos son simétricos respecto a ella. Sin embargo, también es sensible a valores extremos, lo que puede afectar su interpretación en conjuntos de datos con gran variabilidad. Watier et al. (2011) destacan que la media no solo resume los datos, sino que también facilita la comprensión de las distribuciones cuando se combina con otras herramientas estadísticas.

3.2.1. Pasos para Calcular la Media de Datos Simples

Para calcular la media aritmética simple de un conjunto de datos, se siguen estos pasos:

- Sumar todos los valores del conjunto de datos.
- Dividir esta suma por el número total de elementos en el conjunto.

La fórmula matemática para calcular la media aritmética se expresa de la siguiente manera:

$$\bar{x} = \frac{\sum_{i=1}^k x_i}{N}$$

Donde:

N=número total de datos.

x_i =valores de cada dato de la muestra.

3.2.2. Pasos para Calcular la Media de Datos Agrupados

Calcular el Punto Medio o Marca de Clase

La marca de clase se calcula sumando el límite inferior y el límite superior del intervalo y dividiendo entre dos.

$$MC_i = \frac{L_i + L_s}{2}$$

Donde

L_i es el límite inferior del intervalo

L_s es el límite superior del intervalo.

Multiplicar el Punto Medio por la Frecuencia Correspondiente

Para cada intervalo, multiplica la Marca de Clase MC_i del intervalo por la frecuencia absoluta (fa).

Sumar Todos los Productos de los Puntos Medios y sus Frecuencias

Suma todos los productos obtenidos en el paso anterior.

Calcular la Frecuencia Total

Suma todas las frecuencias para obtener la frecuencia total (N).

Dividir la Suma de los Productos por la Frecuencia Total.

Divide la suma de los productos de los puntos medios por la frecuencia total para obtener la media.

Fórmula para Calcular la Media.

La fórmula general para calcular la media de datos agrupados es:

$$\tilde{x} = \frac{\sum_{i=1}^k MC_i * f_i}{N}$$

Donde:

MC_i = Punto medio del intervalo de clase

f_i es la frecuencia del intervalo de clase

k el número total de intervalos de clase.

3.2.3. Ejemplos para calcular la Media Aritmética datos simples

Ejemplo 1

Supongamos que tenemos el siguiente conjunto de datos que representa las edades de un grupo de personas: 25, 30, 35, 40 y 45 años. Sumamos todos los valores:

$$25+30+35+40+45=175$$

Como hay 5 elementos en el conjunto, dividimos la suma (175) por 5 (número total de elementos):

$$\text{Media} = 175/5 = 35.$$

Por lo tanto, la media aritmética de este conjunto de edades es 35 años. Esto indica que, en promedio, la edad de las personas en el grupo es de 35 años.

Ejemplo 2

En una obra de construcción, se han realizado pruebas de resistencia a la compresión en cilindros de concreto. Los resultados en megapascuales (MPa) de las pruebas son los siguientes: 28, 30, 27, 29, 31, 32, 30. ¿Cuál es el promedio de resistencia a la compresión de los cilindros de concreto?

$$\bar{x} = \frac{\sum_1^7 28 + 30 + 27 + 29 + 31 + 32 + 30}{7}$$

$$\bar{x} = \frac{207}{7}$$

$$\bar{x} = 29.57 \text{ MPa}$$

Ejemplo 3

Se ha medido el caudal de agua en un canal durante una semana, y los resultados en metros cúbicos por segundo (m^3/s) son: 2.5, 3.0, 2.8, 2.7, 2.9, 3.1 ¿Cuál es el caudal promedio del canal durante la semana?

$$\bar{x} = \frac{\sum_1^6 2.5 + 3.0 + 2.8 + 2.7 + 2.9 + 3.1}{6}$$

$$\bar{x} = \frac{17}{6}$$

$$\bar{x} = 2.83 \text{ m}^3/\text{s}$$

Ejemplo 4

Durante la construcción de una carretera, se compraron los siguientes lotes de asfalto a distintos precios por tonelada: \$150, \$160, \$155, \$165, \$158, \$162, \$157, 162, \$157, 162, \$157 ¿Cuál es el costo promedio por tonelada de asfalto?

$$\bar{x} = \frac{\sum_1^{11} 150 + 160 + 155 + 165 + 158 + 162 + 157 + 162 + 157 + 162 + 157}{11}$$

$$\bar{x} = \frac{1745}{11}$$

$$\bar{x} = 158.64 \text{ dólares}$$

3.2.4. Ejemplos para calcular la Media Aritmética datos agrupados

Ejemplo 1

Se realizaron pruebas de resistencia a la compresión en bloques de concreto. Los resultados se agruparon en la tabla de frecuencias (ver Tabla 12).

Tabla 12

Frecuencia de resistencia de compresión al concreto

Intervalo de Resistencia (MPa)	Frecuencia (fa)
20-24	3
25-29	7
30-34	15
35-39	5
Total	30

Paso 1: Calcular la marca de clase de cada intervalo:

$$MC_1 = \frac{20+24}{2} = 22 \quad MC_2 = \frac{25+29}{2} = 27 \quad MC_3 = \frac{30+34}{2} = 32 \quad MC_4 = \frac{35+39}{2} = 37$$

Paso 2: Multiplicar cada Marca de clase por su frecuencia, como se muestra en la Tabla 13.

Tabla 13

Frecuencia de resistencia de compresión al concreto

Intervalo de Resistencia (MPa)	Frecuencia (fa)	Marca de clase	Mc*fa
20-24	3	22	66
25-29	7	27	189
30-34	15	32	480
35-39	5	37	185
Total	30		920

Paso 3: Dividir la suma de la marca de clase entre el total de observaciones.

$$\bar{x} = \frac{920}{30} = 30.67 \text{ MPa}$$

Resultado: La media aritmética de la resistencia a la compresión es 30.67 MPa.

Ejemplo 2:

Se tiene en la Tabla 14 con las alturas de edificios en un área determinada, agrupadas en intervalos.

Tabla 14*Alturas de edificios*

Intervalo de Altura (m)	Frecuencia (fa)
30-39	4
40-49	8
50-59	6
60-69	2
Total	20

Paso 1: Calcular la marca de clase de cada intervalo.

$$MC_1 = \frac{30+39}{2} = 34.5 \quad MC_2 = \frac{40+49}{2} = 44.5$$

$$MC_3 = \frac{50+59}{2} = 54.5 \quad MC_4 = \frac{60+69}{2} = 64.5$$

Paso 2: Multiplicar cada Marca de clase por su frecuencia, como se muestra en la Tabla 15.

Tabla 15*Alturas de edificios*

Intervalo de Altura (m)	Frecuencia (fa)	Marca de clase	Mc*fa
30-39	4	34,5	138
40-49	8	44,5	356
50-59	6	54,5	327
60-69	2	64,5	129
Total	20		950

Paso 3: Dividir la suma de la marca de clase entre el total de observaciones.

$$\bar{x} = \frac{950}{20} = 47.5 \text{ m}$$

Resultado: La media aritmética de las alturas de los edificios es 47.5 m.

Ejemplo 3:

Se han medido los espesores de capas de asfalto en diferentes tramos de una carretera, y los datos se agrupan en la Tabla 16.

Tabla 16

Espesor de Capas de Asfalto en una Carretera

Intervalo de Espesor (cm)	Frecuencia (fa)
5,0-5,9	10
6,0-6,9	15
7,0-7,9	20
8,0-8,9	5
Total	50

Paso 1: Calcular la marca de clase de cada intervalo.

$$MC_1 = \frac{5.0+5.9}{2} = 5.45 \quad MC_2 = \frac{6.0+6.9}{2} = 6.45$$

$$MC_3 = \frac{7.0+7.9}{2} = 7.45 \quad MC_4 = \frac{8.0+8.9}{2} = 8.45$$

Paso 2: Multiplicar cada Marca de clase por su frecuencia, como se muestra en la Tabla 17.

Tabla 17

Espesor de Capas de Asfalto en una Carretera

Intervalo de Espesor (cm)	Frecuencia (fa)	Marca de clase	Mc*fa
5,0-5,9	10	5,45	54,5
6,0-6,9	15	6,45	96,75
7,0-7,9	20	7,45	149
8,0-8,9	5	8,45	42,25
Total	50		342,5

Paso 3: Dividir la suma de la marca de clase entre el total de observaciones.

$$\bar{x} = \frac{342.5}{50} = 6.85 \text{ cm}$$

Resultado: La media aritmética del espesor del asfalto es 6.85 cm.

3.3. Mediana

La mediana es una medida estadística que representa el valor central en un conjunto de datos ordenados de menor a mayor (o de mayor a menor). A diferencia de la media, la mediana no se ve afectada por valores extremos o atípicos, lo que la hace adecuada para describir distribuciones asimétricas.

Esta medida de tendencia central es especialmente útil en conjuntos de datos donde los valores extremos podrían distorsionar la interpretación de la media. Según McGreevy et al. (2009), el uso de la mediana como indicador central es preferible en datos muy sesgados o no simétricos, ya que proporciona una representación más robusta y confiable de la ubicación central de los datos.

Según McGreevy et al. (2009), el uso de la mediana como indicador central es preferible en datos muy sesgados o no simétricos, ya que proporciona una representación más robusta y confiable de la ubicación central de los datos.

3.3.1. Pasos para Calcular la Mediana de Datos Simples

Para calcular la mediana de un conjunto de datos, se siguen estos pasos:

- Ordenar los datos de menor a mayor (o de mayor a menor).
- Identificar el valor que ocupa la posición central en la lista ordenada.

Si el conjunto de datos tiene un número impar de elementos, la mediana será el valor que está en la posición central. Si el conjunto de datos tiene un número par de elementos, la mediana será el promedio de los dos valores centrales.

x
$x_{(1)}$
$x_{(2)}$
$\vdots$
$x_{(n)}$

$$\tilde{x} = x_{(n+1)/2}$$

Si n es impar

$$\tilde{x} = \frac{x_{(\frac{n}{2})} + x_{(\frac{n}{2}+1)}}{2}$$

Si n es par

3.3.2. Pasos para Calcular la Mediana de Datos Agrupados

Calcular la mediana de datos agrupados en intervalos requiere seguir una serie de pasos específicos. A continuación, se describen estos pasos:

- **Organizar la Tabla de Frecuencias:**

Asegúrate de que los datos estén organizados en una tabla de frecuencias que incluya intervalos de clase, frecuencias (f_a), y la frecuencia acumulada (Fa).

- **Determinar el Intervalo de la Mediana:**

Encuentra el intervalo de la mediana, que es el intervalo donde se encuentra la mediana. Para ello, localiza el intervalo donde la frecuencia acumulada es mayor o igual que la mitad del total de datos, es decir: $\frac{N}{2}$

- **Aplicar la Fórmula de la Mediana:**

Una vez identificado el intervalo de la mediana, utiliza la siguiente fórmula para calcularla:

$$\tilde{X} = L_I + \left(\frac{\frac{N}{2} - Fa}{fa} \right) * A$$

Donde:

L_I = Límite inferior del intervalo de la mediana.

N = Total de datos.

Fa = Frecuencia acumulada anterior al intervalo de la mediana.

fa = Frecuencia del intervalo de la mediana.

A = Amplitud del intervalo (tamaño de clase).

3.3.3. Ejemplo de Mediana

Ejemplo 1:

Supongamos que tenemos el siguiente conjunto de datos que representa el número de horas que un grupo de estudiantes estudia cada día: 4, 5, 6, 7, 8.

Solución:

Ordenamos los datos de menor a mayor: 4, 5, 6, 7, 8.

Como el conjunto tiene un número impar de elementos (5 en total), la mediana será el valor que ocupa la posición central, que es el tercer valor en la lista ordenada, que es 6.

Por lo tanto, la mediana de este conjunto de datos es 6 horas. Esto indica que el valor que está en la posición central de las horas de estudio es 6, lo que sugiere que la mitad de los estudiantes estudian más de 6 horas y la otra mitad estudian menos de 6 horas.

La mediana es una medida robusta de tendencia central que es especialmente útil cuando los datos tienen valores extremos o cuando se busca una representación de la ubicación central que no se vea afectada por valores atípicos en el conjunto de datos.

Ejemplo 2:

Se han registrado los tiempos de construcción (en semanas) de varios proyectos de edificios residenciales: 52, 48, 50, 47, 49, 51, 46, 53. ¿Cuál es la mediana del tiempo de construcción

Solución:

Ordenar los valores de menor a mayor: 46, 47, 48, 49, 50, 51, 52, 53.

Como hay un número par de valores (8), la mediana es el promedio de los dos valores centrales (cuarto y quinto):

$$\tilde{x} = \frac{49+50}{2}$$

$$\tilde{x} = 49.5 \text{ semanas}$$

Ejemplo 3:

En un estudio de planificación urbana, se registraron las alturas de varios edificios en metros: 55, 60, 58, 65, 62, 57, 59. ¿Cuál es la mediana de las alturas de estos edificios?

Solución:

Ordenar los valores de menor a mayor: 55, 57, 58, 59, 60, 62, 65.

Encontrar el valor medio (en este caso, el cuarto valor porque hay 7 datos):

$$\tilde{x} = 59 \text{ metros}$$

Ejemplo 4:

Durante la instalación de un sistema de alcantarillado, se midieron los diámetros (en cm) de varios tubos seleccionados al azar: 10, 15, 12, 18, 14, 16, 13. ¿Cuál es la mediana del diámetro de estos tubos?

Solución:

Ordenar los valores de menor a mayor: 10, 12, 13, 14, 15, 16, 18.
Encontrar el valor medio (en este caso, el cuarto valor porque hay 7 datos):

$$\tilde{x} = 14 \text{ cm.}$$

Ejemplo 5:

Supongamos que tienes en la tabla 17 de frecuencias con intervalos de resistencia a la compresión de concreto. Calcule la mediana.

Tabla 17

Resistencia a la compresión de concreto

Intervalo de Resistencia (MPa)	Frecuencia (fa)	Frecuencia Acumulada (Fa)
20-24	3	3
25-29	7	10
30-34	15	25
35-39	5	30
Total	50	

Paso 1. Calcula

$$\frac{N}{2} = \frac{30}{2} = 15$$

Paso 2. Encuentra el intervalo de la mediana.

El intervalo de la mediana es 30-34, ya que $Fa = 10$ y $Fa = 25$ lo que incluye a 15.

Paso 3: Aplica la fórmula:

$$L_I = 30 \text{ MPa.}$$

$$N = 30.$$

$$Fa = 10.$$

$$fa = 15$$

$$A = 5$$

$$\tilde{X} = 30 + \left(\frac{\frac{30}{2} - 10}{15} \right) * 5 = 31.67 \text{ MPa}$$

Resultado: La mediana de la resistencia a la compresión es 31.67 MPa.

3.4. Moda

La moda es una medida estadística que identifica el valor que aparece con mayor frecuencia en un conjunto de datos. Este indicador es útil para describir el valor más común o frecuente en un conjunto de observaciones y puede ser aplicado en datos con distribuciones unimodales o multimodales.

Para determinar la moda de un conjunto de datos, se deben contar las frecuencias de aparición de cada valor y encontrar aquel o aquellos con la mayor frecuencia. Es posible que un conjunto de datos sea unimodal (una sola moda), multimodal (más de una moda) o que no tenga moda si todos los valores tienen la misma frecuencia. Según Jankowski y Flannelly (2015), la moda es particularmente útil en investigaciones que emplean escalas nominales, ya que proporciona una forma sencilla de identificar categorías predominantes.

3.4.1. Ejemplo de Moda

Ejemplo 1:

Supongamos que tenemos el siguiente conjunto de datos que representa las calificaciones de un grupo de estudiantes en un examen: 75, 80, 85, 80, 90, 80, 85.

Identificamos cuántas veces aparece cada valor en el conjunto de datos:

75 aparece una vez.

80 aparece tres veces.

85 aparece dos veces.

90 aparece una vez.

El valor que tiene la mayor frecuencia de aparición es 80, ya que aparece tres veces en el conjunto de datos.

Por lo tanto, la moda de este conjunto de datos es 80. Esto indica que la calificación más común entre los estudiantes fue 80 puntos en el examen.

La moda es una medida simple pero efectiva de tendencia central que nos permite identificar rápidamente el valor más frecuente en un conjunto de datos, lo que puede ser útil para entender la distribución y características principales de los datos.

Ejemplo 2:

En una construcción, se utilizan diferentes diámetros de barras de refuerzo de acero. Los diámetros en milímetros (mm) de las barras utilizadas en una sección son: 10, 12, 10, 16, 10, 12, 16, 12. ¿Cuál es la moda de los diámetros de las barras de refuerzo?

Solución:

Contar la frecuencia de cada diámetro:

10 mm: 3 veces

12 mm: 3 veces

16 mm: 2 veces

Determinar la moda (el valor que más se repite):

La moda es 10 mm y 12 mm (bimodal).

Moda = 10 mm y 12 mm

Ejemplo 3:

En un complejo residencial, se han registrado las alturas de techos en metros (m) de varios apartamentos: 2.5, 2.8, 2.5, 2.6, 2.7, 2.5, 2.8, 2.8. ¿Cuál es la moda de las alturas de los techos?

Solución:

Contar la frecuencia de cada altura:

2.5 m: 3 veces

2.6 m: 1 vez

2.7 m: 1 vez
2.8 m: 3 veces

Determinar la moda (el valor que más se repite):
La moda es 2.5 m y 2.8 m (bimodal).
Moda=2.5m y 2.8m

Ejemplo 4:

En una obra de construcción, se han identificado varios tipos de suelo: arcilla, arena, grava, arcilla, arena, arcilla, arena, grava, arena. ¿Cuál es la moda de los tipos de suelo?

Solución:

Contar la frecuencia de cada tipo de suelo:
Arcilla: 3 veces
Arena: 4 veces
Grava: 2 veces

Determinar la moda (el valor que más se repite):
Moda=arena

Ejemplo 5:

Se ha recopilado la cantidad de pisos de varios edificios comerciales: 4, 5, 4, 6, 4, 5, 6, 5, 4. ¿Cuál es la moda de la cantidad de pisos en estos edificios comerciales?

Solución:

Contar la frecuencia de cada cantidad de pisos:
4 pisos: 4 veces
5 pisos: 3 veces
6 pisos: 2 veces

Determinar la moda (el valor que más se repite):
Moda=4 pisos

3.5. Usos y aplicaciones de las medidas de tendencia central

Las medidas de tendencia central, como la media, mediana y moda, son herramientas fundamentales en el campo de la ingeniería civil y arquitectura. Estas medidas permiten analizar y comprender mejor los datos recopilados en diversas áreas, desde el diseño de estructuras hasta la gestión de proyectos de construcción. A continuación, exploraremos en detalle los usos y aplicaciones de estas medidas en la ingeniería civil.

3.5.1. Uso de la Media en Ingeniería Civil y Arquitectura

La media aritmética es una medida de tendencia central fundamental en la ingeniería civil y la arquitectura, ya que proporciona un valor representativo del conjunto de datos analizados. Este valor es utilizado en múltiples aplicaciones que permiten mejorar el diseño, la gestión y el análisis en estos campos. A continuación, se presentan ejemplos concretos de su uso:

- 1. Diseño de Estructuras:** En el diseño de puentes y edificios, la media se utiliza para calcular cargas promedio, resistencia de materiales y distribución de cargas. Esto permite a los ingenieros garantizar que las estructuras sean seguras y funcionales. Según Sergeeva (2020), cualquier diseño estructural debe estar matemáticamente justificado, y la media es una herramienta clave para procesar datos experimentales y modelar comportamientos estructurales.
- 2. Análisis de Rendimiento:** La media es utilizada en proyectos de construcción para evaluar el rendimiento promedio de equipos y trabajadores, así como el consumo de recursos. Esto ayuda a identificar ineficiencias y a optimizar procesos.
- 3. Estimación de Costos:** En la planificación financiera, la media se aplica para calcular costos promedio considerando precios de materiales, tarifas laborales y otros factores. Esto proporciona una base confiable para presupuestos iniciales y ajustes durante la ejecución del proyecto.
- 4. Control de Calidad:** En el control de calidad, la media es esencial para evaluar la resistencia y durabilidad de materiales como

concreto y acero. Parkhurst (1998) señala que la media aritmética es preferida por su facilidad de cálculo y su relevancia científica en la representación de datos de concentración o propiedades materiales.

5. **Análisis de Resistencia de Materiales:** La media de resultados obtenidos en pruebas de compresión y tracción se utiliza para determinar la calidad promedio de los materiales, asegurando que cumplan con los estándares requeridos.
6. **Planificación de Proyectos:** En la estimación de tiempos de construcción, se utiliza la media de duraciones de proyectos similares para prever cronogramas más precisos.

3.5.2. Uso de la Mediana en Ingeniería Civil y Arquitectura

La mediana es una medida estadística de tendencia central que es particularmente valiosa en situaciones donde los datos contienen valores atípicos o distribuciones sesgadas. A continuación, se describen algunas aplicaciones relevantes de la mediana en los campos de la ingeniería civil y la arquitectura:

- **Diseño Vial:** En el diseño de carreteras y vías de transporte, la mediana se utiliza para determinar la ubicación central de puntos críticos como curvas pronunciadas o zonas de pendientes. Según Wesolowsky (1975), la mediana lineal es una herramienta práctica para optimizar el diseño de corredores de transporte minimizando las distancias promedio desde puntos de interés.
- **Evaluación de Riesgos Geotécnicos:** En proyectos de ingeniería civil, la mediana se utiliza para identificar valores representativos en datos geotécnicos como resistencia del suelo y permeabilidad. Esto permite obtener estimaciones más robustas en condiciones de datos extremos o variables. Según Minsker y Strawn (2023), la mediana geométrica proporciona una mejor representación central en distribuciones con colas pesadas, lo que es ideal para análisis de suelos.
- **Planificación de Recursos:** En la planificación de recursos como maquinaria y equipos, la mediana se utiliza para determinar el tiempo típico de vida útil o intervalos de mantenimiento, evitando

que valores extremos distorsionen las estimaciones. Este enfoque asegura una asignación más precisa de recursos en proyectos de construcción.

- **Distribución de Espacios:** En estudios de diseño arquitectónico, la mediana de áreas de habitaciones puede usarse para representar el tamaño típico de los espacios en un edificio, especialmente cuando hay gran variabilidad en las dimensiones. Levy (2008) señala que la mediana es útil en física y astronomía para analizar datos contaminados por valores extremos, un principio que también se aplica en contextos arquitectónicos.
- **Análisis de Datos de Encuestas:** En encuestas sobre preferencias de diseño o satisfacción de usuarios, la mediana identifica la respuesta típica sin ser influenciada por valores extremos. Esto es particularmente útil en estudios de percepción de usuarios sobre diseño urbano o arquitectónico.

3.5.3. Uso de la Moda en Ingeniería Civil y Arquitectura

La moda, como medida estadística, permite identificar el valor más frecuente dentro de un conjunto de datos, lo que resulta valioso para analizar patrones y tendencias en diversas áreas de la ingeniería civil y la arquitectura. Algunos de sus usos principales incluyen:

- **Análisis de Tráfico:** En el diseño de sistemas de transporte y análisis de tráfico, la moda se utiliza para identificar patrones de flujo vehicular en diferentes horarios y áreas geográficas. Esto ayuda a optimizar las infraestructuras viales y a gestionar de manera más eficiente la circulación vehicular. Según Florian et al. (1979), el análisis de modos de transporte permite integrar modelos de tráfico para evaluar patrones en sistemas urbanos y rurales.
- **Gestión de Inventarios:** En el manejo de materiales de construcción, la moda se utiliza para identificar los productos más demandados y optimizar los niveles de inventario. Este análisis ayuda a reducir costos al mantener existencias adecuadas y evitar desabastecimientos.

- **Planificación Urbana:** La moda se emplea para identificar los usos del suelo más comunes en zonas urbanas (residenciales, comerciales, industriales), facilitando decisiones informadas en el desarrollo y diseño urbano. Zheng et al. (2011) explican cómo los modelos de planeación urbana utilizan patrones modales para coordinar el uso del suelo y el transporte, integrando de manera eficiente los servicios urbanos.

Ejercicios a desarrollar del Capítulo 3

1. Se han medido las longitudes de cinco vigas de acero utilizadas en una construcción. Las longitudes, en metros, son: 4.2, 3.8, 4.0, 4.1, y 3.9. Determine la longitud media de las vigas.
2. Un arquitecto ha registrado los costos de cinco tipos diferentes de materiales, en miles de dólares: 12, 15, 13, 14, y 16. Calcule el costo medio de los materiales.
3. Se han recopilado las alturas de seis edificios recientemente diseñados por un estudio de arquitectura, en metros: 50, 55, 60, 65, 70, y 75. Determine la altura media de estos edificios.
4. Un ingeniero civil ha registrado las duraciones, en meses, de siete proyectos diferentes: 12, 14, 16, 18, 20, 22, y 24. Calcule la duración media de los proyectos.
5. En ocho proyectos de construcción, se registró el consumo de hormigón, en metros cúbicos: 320, 340, 310, 330, 350, 360, 340, y 320. Determine la cantidad media de hormigón utilizado por proyecto.
6. Se han estimado los costos de nueve proyectos de ingeniería civil, en miles de dólares: 200, 220, 210, 230, 240, 250, 210, 220, y 230. Calcule el costo medio de estos proyectos.
7. Se han medido las alturas de varias columnas en un proyecto de construcción. Los datos recopilados se presentan en la Tabla 18. Calcule la altura media de las columnas.
8. En un diseño arquitectónico, se han registrado las áreas de varios espacios comunes. Los datos se presentan En la Tabla 19. Determine el área media de los espacios comunes.
9. Se han estimado los costos de varios proyectos de construcción. Los datos se presentan en la Tabla 20. Calcule el costo medio de los proyectos.
10. La duración de varias obras se ha registrado en la Tabla 21. Determine la duración media de las obras
11. Se han probado diferentes pilares para determinar la carga máxima que pueden soportar. Los resultados se presentan en la

- Tabla 22. Calcule la carga media soportada por los pilares.
12. En varios proyectos de construcción, se ha registrado el volumen de hormigón utilizado. Los datos se presentan en la Tabla 23. Determine el volumen medio de hormigón utilizado por proyecto.
 13. En un proyecto de construcción, se registran los espesores de muros de concreto en diferentes secciones de un edificio. Los datos se agrupan en la Tabla 24. Calcule la mediana del espesor de los muros.
 14. Se han medido las áreas de las fachadas de varios edificios en un desarrollo urbano y se agrupan en la Tabla 25. Calcule la mediana del espesor de los muros.
 15. Se han medido las áreas de las fachadas de varios edificios en un desarrollo urbano y se agrupan en la Tabla 26. Calcule la mediana del área de las fachadas.
 16. Se ha realizado un estudio sobre la cantidad de ventanas en edificios residenciales, agrupadas en la Tabla 27. Calcule la mediana de la cantidad de ventanas en los edificios residenciales.
 17. El tiempo que tomó completar varios proyectos arquitectónicos se ha registrado y agrupado en la Tabla 28. Calcule la mediana del tiempo de construcción de estos proyectos arquitectónicos.

Tabla 18

Altura de columnas

Altura (m)	Frecuencia
2.5	3
3.0	5

Tabla 19

Área de espacios comunes

Área (m ²)	Frecuencia
50	4
75	6
100	2

Tabla 20

Costos de proyectos

Costo (mil \$)	Frecuencia
200	3
250	5
300	2

Tabla 21

Duración de obras

Duración (meses)	Frecuencia
6	4
8	6
10	3

Tabla 22*Duración de obras*

Carga (toneladas)	Frecuencia
10	5
15	8
20	4

Tabla 23*Volumen de Hormigón*

Volumen (m³)	Frecuencia
100	10
150	15
200	5

Tabla 24*Espesor de muros*

Intervalo de Espesor (cm)	Frecuencia (fa)
10-14	5
15-19	12
20-24	18

Tabla 25*Áreas de fachadas*

Intervalo de Área (m²)	Frecuencia (fa)
100-149	8
150-199	15
200-249	20
250-299	7

Tabla 26*Altura de Pisos en un Edificio*

Intervalo de Altura (m)	Frecuencia (fa)
3,0-3,4	10
3,5-3,9	18
4,0-4,4	12
4,5-4,9	5

Tabla 27*Cantidad de Ventanas en Edificios Residenciales*

Intervalo de Ventanas	Frecuencia (fa)
10-14	6
15-19	14
20-24	11
25-29	4

Tabla 28*Tiempo de Construcción de Proyectos Arquitectónicos*

Intervalo de Tiempo (meses)	Frecuencia (fa)
12-15	7
16-19	13
20-23	16
24-27	9

18. Se han medido las alturas (en metros) de cinco edificios en un distrito urbano: 120, 150, 130, 140, 135. Determine la mediana de las alturas de los edificios.
19. Se han registrado las áreas (en hectáreas) de seis parques urbanos: 2.5, 3.0, 3.5, 4.0, 4.5, 5.0. Calcule la mediana del área de los parques.
20. Se han estimado los costos (en miles de dólares) de ocho materiales de construcción: 20, 22, 19, 21, 23, 24, 18, 25.

Determine la mediana de los costos de los materiales.

21. La duración (en meses) de siete proyectos de construcción son: 12, 16, 14, 18, 15, 17, 13. Calcule la mediana de la duración de los proyectos.
22. En diez proyectos de construcción, se ha registrado el volumen de hormigón utilizado (en metros cúbicos): 320, 340, 310, 330, 350, 360, 340, 320, 330, 310. Determine la mediana del volumen de hormigón utilizado.
23. Se han medido las alturas (en metros) de nueve puentes: 50, 55, 60, 65, 70, 75, 80, 85, 90. Calcule la mediana de las alturas de los puentes.
24. Se han registrado las alturas (en metros) de diferentes edificaciones en un área urbana. Los datos son: 100, 120, 130, ?, 150, 160, donde “?” representa una altura desconocida. Suponiendo que la altura desconocida no afecta el orden de los datos, determine la mediana.
25. Las áreas (en hectáreas) de varios lotes son: <5, 6, 7, ?, >10. Asumiendo que el lote con área desconocida se encuentra en el rango medio, calcule la mediana.
26. Los costos (en miles de dólares) de varios proyectos son: <200, 220, 240, ?, 280, >300. Si el proyecto con costo desconocido tiene un costo mayor que 240 pero menor que 280, determine la mediana.
27. Las duraciones (en meses) de varios proyectos de construcción son: <12, 14, 16, ?, 20, >22. Si se sabe que el proyecto con duración desconocida tardó más de 16 meses, pero menos de 20, encuentre la mediana.
28. Las cargas máximas (en toneladas) soportadas por diferentes pilares son: <15, 17, 19, ?, 23, 25, >27. Considerando que el pilar con carga desconocida soporta más que 19 pero menos que 23 toneladas, calcule la mediana.
29. Los volúmenes (en metros cúbicos) de hormigón utilizados en varios proyectos son: <300, 320, 340, ?, 380, 400, >420. Si el volumen desconocido es mayor que 340 pero menor que 380 metros cúbicos, determine la mediana.

- 30.** Se han medido las alturas (en metros) de varios edificios, resultando en la siguiente distribución de frecuencias:
- 100-120m: 3 edificios
 - 120-140m: 5 edificios
 - 140-160m: 2 edificios
- Determine la mediana de las alturas de los edificios.
- 31.** Las áreas (en hectáreas) de varios terrenos están distribuidas de la siguiente manera:
- 1-3 ha: 4 terrenos
 - 3-5 ha: 6 terrenos
 - 5-7 ha: 2 terrenos
- Calcule la mediana del área de estos terrenos.
- 32.** La duración (en meses) de varios proyectos de construcción se distribuye así:
- 6-12 meses: 3 proyectos
 - 12-18 meses: 7 proyectos
 - 18-24 meses: 5 proyectos
- Determine la mediana de la duración de estos proyectos.
- 33.** Los costos (en miles de dólares) de materiales de construcción se distribuyen de la siguiente manera:
- 10-20k: 4 materiales
 - 20-30k: 8 materiales
 - 30-40k: 3 materiales
- Calcule la mediana de los costos de estos materiales.
- 34.** Los volúmenes (en metros cúbicos) de hormigón utilizados en varios proyectos se distribuyen así:
- 100-200 m³: 5 proyectos
 - 200-300 m³: 10 proyectos
 - 300-400 m³: 5 proyectos
- Determine la mediana del volumen de hormigón utilizado.
- 35.** Las alturas (en metros) de varias torres de comunicación se distribuyen de la siguiente manera:
- 30-50 m: 4 torres
 - 50-70 m: 8 torres
 - 70-90 m: 3 torres

Calcule la mediana de las alturas de estas torres.

36. En un proyecto de construcción, se han utilizado los siguientes materiales: hormigón, acero, madera, acero, hormigón, madera, acero. Determine la moda de los materiales utilizados.
37. En una revisión de proyectos arquitectónicos, se observaron los siguientes estilos de diseño: moderno, clásico, moderno, minimalista, clásico, moderno. Identifique la moda en los estilos de diseño.
38. En una urbanización, se han utilizado los siguientes tipos de pavimentos: asfalto, adoquín, asfalto, concreto, adoquín, adoquín, concreto. Determine la moda de los tipos de pavimentos.
39. En un barrio residencial, los colores de las fachadas de las casas son: rojo, azul, verde, azul, verde, verde, rojo. Calcule la moda de los colores de las fachadas.
40. En un conjunto de oficinas, los tipos de iluminación son: LED, fluorescente, incandescente, LED, LED, fluorescente. Identifique la moda en los tipos de iluminación.
41. En una serie de edificios comerciales, los estilos de ventanas son: panorámico, clásico, panorámico, clásico, panorámico, abatible. Calcule la moda de los estilos de ventanas.
42. Se midieron las alturas (en metros) de varios edificios en una ciudad: 120, 150, 130, 120, 140, 150, 120. Determine la moda de las alturas de los edificios.
43. Las longitudes (en metros) de una serie de puentes son: 200, 250, 200, 300, 250, 200. Identifique la moda en las longitudes de los puentes.
44. Se registraron las áreas (en hectáreas) de varios parques urbanos: 2.5, 3.0, 2.5, 4.0, 3.5, 2.5, 3.0. Calcule la moda de las áreas de los parques.
45. Los caudales (en metros cúbicos por segundo) de varios ríos fueron: 80, 120, 100, 80, 120, 80. Determine la moda de los caudales de los ríos.
46. Se contabilizó la cantidad de pisos de varios edificios: 10, 15, 10, 20, 15, 10, 15. Identifique la moda en la cantidad de pisos

de los edificios.

- 47.** Los anchos (en metros) de diferentes tramos de carreteras son: 12, 14, 16, 12, 18, 14, 12. Calcule la moda de los anchos de las carreteras.
- 48.** Se han agrupado las alturas (en metros) de varios edificios: [100-150], [150-200], [200-250], con frecuencias de 5, 8 y 3 respectivamente. Determine la moda de las alturas de los edificios.
- 49.** Las longitudes (en metros) de una serie de puentes se agruparon en: [100-200], [200-300], [300-400], con frecuencias de 4, 7 y 2 respectivamente. Identifique la moda en las longitudes de los puentes.
- 50.** Se registraron las áreas (en hectáreas) de varios parques urbanos en intervalos: [1-3], [3-5], [5-7], con frecuencias de 6, 9 y 4 respectivamente. Calcule la moda de las áreas de los parques.
- 51.** Los caudales (en metros cúbicos por segundo) de varios ríos se agruparon en: [50-100], [100-150], [150-200], con 5, 10 y 3 registros respectivamente. Determine la moda de los caudales de los ríos.
- 52.** Se contabilizó la cantidad de pisos de varios edificios en intervalos: [0-10], [10-20], [20-30], con 8, 12 y 5 edificios respectivamente. Identifique la moda en la cantidad de pisos de los edificios.
- 53.** Los anchos (en metros) de diferentes tramos de carreteras se agruparon en: [10-15], [15-20], [20-25], con 7, 11 y 4 tramos respectivamente. Calcule la moda de los anchos de las carreteras.

Conclusiones del Capítulo 3

El capítulo tres permitió comprender el papel esencial de las medidas de tendencia central en el análisis y la interpretación de datos en ingeniería civil y arquitectura. Estas medidas, que incluyen la media aritmética, la mediana y la moda, facilitaron la identificación del valor central de un conjunto de datos, proporcionando información clave para la toma de decisiones fundamentadas en evidencia. Su correcta aplicación resultó indispensable para la planificación, diseño y ejecución de proyectos, donde el manejo de datos precisos optimizó los recursos y mejoró la eficiencia de las construcciones.

La media aritmética demostró ser una herramienta valiosa para la estimación de costos, tiempos de construcción y propiedades de los materiales. Su capacidad para generar promedios permitió establecer presupuestos realistas y comparar diferentes opciones con base en criterios objetivos. Sin embargo, se evidenció que su sensibilidad a valores extremos podía distorsionar el análisis, lo que destacó la importancia de utilizar medidas complementarias.

La mediana resultó especialmente útil en casos donde los datos presentaron distribuciones sesgadas. En proyectos de construcción, por ejemplo, permitió obtener estimaciones más precisas cuando existieron costos atípicamente altos o bajos que habrían afectado el cálculo del promedio. Esta propiedad de la mediana la convirtió en una alternativa robusta para la interpretación de datos en contextos donde la variabilidad es significativa.

Por otro lado, la moda brindó una perspectiva diferente al identificar los valores más frecuentes dentro de un conjunto de datos. Su aplicación en arquitectura y construcción permitió reconocer tendencias en el uso de materiales, características estructurales más comunes y patrones recurrentes en fallas o defectos. Esta capacidad de detectar frecuencias dominantes resultó valiosa para optimizar diseños y mejorar la calidad de los proyectos.

El análisis de las medidas de tendencia central en este capítulo reafirmó que cada una de ellas posee fortalezas y limitaciones que deben ser consideradas según el contexto del análisis. La

combinación de estas herramientas estadísticas ofreció una visión más completa de los datos, permitiendo realizar comparaciones efectivas y tomar decisiones estratégicas fundamentadas en información confiable.

En definitiva, la correcta aplicación de la media, la mediana y la moda fortaleció la capacidad de los profesionales para interpretar datos y optimizar procesos en la ingeniería civil y la arquitectura. La incorporación de estas herramientas en la planificación y ejecución de proyectos permitió mejorar la precisión en la gestión de costos, tiempos y recursos, asegurando mayor eficiencia y calidad en el desarrollo de infraestructuras.

Capítulo 4

MEDIDAS DE DISPERSIÓN

En este capítulo, exploraremos conceptos clave como la desviación estándar, la varianza, el rango y el rango intercuartílico. Estas medidas nos proporcionan información valiosa sobre la dispersión de nuestros datos; es decir, cuánto se alejan los valores individuales de la medida central representada por la media, mediana o moda.



4.1. Introducción

Comprender la dispersión de los datos es esencial en Arquitectura e la ingeniería civil, donde la variabilidad puede influir en decisiones críticas en el diseño de estructuras, gestión de proyectos y evaluación de riesgos. Por lo tanto, este capítulo nos llevará a explorar aplicaciones prácticas de las medidas de dispersión en diferentes contextos, proporcionando una visión integral de su importancia y utilidad en el análisis de datos.

4.2. Rango

El rango es una medida estadística de dispersión que indica la amplitud total de un conjunto de datos. Es la diferencia entre el valor máximo y el valor mínimo del conjunto, proporcionando una visión básica de la variabilidad en los datos. Este indicador es fácil de calcular y útil para evaluar la extensión de las observaciones, aunque no ofrece detalles sobre cómo se distribuyen los valores dentro del rango.

Para calcular el rango de un conjunto de datos, se siguen estos pasos:

1. Identificar el valor máximo (Máx) y el valor mínimo (Mín) en el conjunto de datos.
2. Calcular la diferencia entre estos valores utilizando la fórmula:

$$\text{Rango} = \text{Valor Máximo} - \text{Valor Mínimo}$$

El rango es una herramienta práctica para evaluaciones rápidas, especialmente en contextos como el control de calidad, donde la identificación de extremos es crítica. Según Manikandan (2011), el rango destaca por su simplicidad, pero es altamente sensible a valores atípicos y no utiliza toda la información contenida en los datos, lo que limita su utilidad en distribuciones complejas.

En investigaciones y aplicaciones prácticas, como las pruebas de control de calidad, el rango puede ser un complemento efectivo a otras medidas de dispersión más sofisticadas, como la desviación estándar o el rango intercuartil.

Ejemplo:

Supongamos que tenemos el siguiente conjunto de datos que representa la altura (en centímetros) de cinco árboles en un parque: 150, 160, 170, 180, 190.

Solución:

Identificamos el valor máximo, que es 190, y el valor mínimo, que es 150.

Calculamos el rango restando el valor máximo menos el valor mínimo:

$$\text{Rango} = 190 - 150$$

$$\text{Rango} = 40$$

Por lo tanto, el rango de alturas de los árboles es de 40 centímetros. Esto nos indica que la diferencia entre el árbol más alto y el más bajo es de 40 centímetros, lo que representa la variabilidad en la altura de los árboles en el parque.

- **Datos de Temperaturas:** Supongamos que tenemos los siguientes datos de temperaturas máximas (en grados Celsius) registradas durante una semana: 25, 28, 22, 30, 26, 31, 24. Calcula el rango de estas temperaturas.
- **Datos de Velocidades:** Un vehículo registró las siguientes velocidades (en kilómetros por hora) durante un trayecto: 80, 100, 90, 110, 85. Calcula el rango de estas velocidades.
- **Datos de Edades:** En un estudio sobre la edad de estudiantes de una escuela, se encontraron las siguientes edades (en años): 18, 20, 17, 22, 19, 21. Calcula el rango de estas edades.

Estos ejercicios te permitirán practicar el cálculo del rango y comprender cómo esta medida de dispersión proporciona información sobre la variabilidad de datos en diferentes contextos.

4.3. Varianza

La varianza mide qué tan lejos están los valores individuales del promedio del conjunto de datos. Matemáticamente, se calcula como el promedio de los cuadrados de las diferencias entre cada valor y la

media. Según Fisan et al. (2014), la varianza es ampliamente aceptada como la medida más robusta de dispersión debido a su capacidad para incorporar toda la información del conjunto de datos y ser manipulada algebraicamente en análisis posteriores.

La varianza es útil para describir cuán consistentes son los datos dentro de un conjunto. Por ejemplo, en mediciones experimentales, una varianza baja indica que los datos están estrechamente agrupados alrededor de la media, mientras que una varianza alta sugiere mayor dispersión. Como indica Hayslett y Murphy (1981), su utilidad radica no solo en su capacidad para medir dispersión, sino también en su aplicabilidad en modelos estadísticos avanzados que requieren mediciones precisas de variabilidad.

Se emplea en diversas aplicaciones, desde la evaluación de calidad en procesos industriales hasta el análisis financiero y los modelos predictivos en ciencias naturales. Según Ye et al. (2018), el concepto de varianza también se utiliza en teorías de medición para interpretar intervalos probabilísticos y evaluar errores en sistemas complejos.

Más específicamente, la varianza es una medida de que tan cerca, o que tan lejos están los diferentes valores de su propia media aritmética.

- Cuando más lejos están las X_i de su propia media aritmética, mayor es la varianza.
- Cuando más cerca estén las X_i a su media menos es la varianza.

S será siempre un valor no negativo, que puede ser igual o distinta de 0. Será 0 solamente cuando $X_i = \bar{X}$

Para comprender la varianza, es importante entender que cada valor en un conjunto de datos puede desviarse de la media. La varianza mide cuánto se desvían estos valores individuales de la media, de manera que valores más altos de varianza indican una mayor dispersión de los datos.

La fórmula matemática para calcular la varianza se expresa como:

$$s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}$$

Donde:

x_i = cada valor de la muestra

$\bar{x}$ = media de la muestra

n = tamaño de la muestra

En esta fórmula, se calcula la diferencia entre cada valor individual y la media, se eleva al cuadrado esta diferencia para evitar que las diferencias positivas y negativas se anulen entre sí, se suman todas estas diferencias al cuadrado y se divide entre el número total de elementos en el conjunto.

Ejemplo:

Supongamos que tenemos el siguiente conjunto de datos que representa las edades (en años) de cinco personas: 30, 35, 40, 45, 50.

Calculamos la media de estos datos:

$$\bar{x} = \frac{30 + 35 + 40 + 45 + 50}{5}$$

$$\bar{x} = 40$$

Calculamos la varianza utilizando la fórmula:

$$s^2 = \frac{\sum_{i=1}^5 (30-40)^2 + (35-40)^2 + (40-40)^2 + (45-40)^2 + (50-40)^2}{5-1}$$

$$s^2 = \frac{250}{4}$$

$$s^2 = 62.50$$

Por lo tanto, la varianza de estas edades es 62.50. Esto indica que las edades en este conjunto están dispersas alrededor de su media, con una variabilidad de aproximadamente 62 unidades al cuadrado.

Los siguientes ejercicios te ayudarán a practicar el cálculo de la varianza y a comprender cómo esta medida de dispersión nos

proporciona información valiosa sobre la variabilidad de los datos en diferentes contextos y situaciones.

- **Datos de Alturas:** Considera el siguiente conjunto de datos que representa las alturas (en centímetros) de una muestra de estudiantes: 160, 165, 170, 175, 180. Calcula la varianza de estas alturas.
- **Datos de Tiempos:** Se registraron los siguientes tiempos (en minutos) que tomó completar una actividad para diferentes personas: 10, 12, 15, 9, 11. Calcula la varianza de estos tiempos.
- **Datos de Temperaturas:** Se registraron las siguientes temperaturas (en grados Celsius) durante una semana: 20, 22, 18, 25, 21. Calcula la varianza de estas temperaturas.

4.4. Desviación estándar

La desviación estándar es una medida estadística ampliamente utilizada para evaluar la dispersión de un conjunto de datos respecto a su media. Específicamente, proporciona información sobre cuánto varían los datos individuales en relación con el valor promedio, siendo una herramienta esencial para interpretar la consistencia y variabilidad de las observaciones.

La desviación estándar se calcula como la raíz cuadrada de la varianza, lo que permite expresar la dispersión en las mismas unidades que los datos originales, facilitando su interpretación en contextos prácticos. Según Harrisson (2018), esta métrica es preferida en áreas técnicas porque ofrece una representación clara y directa de la variabilidad dentro de distribuciones complejas.

Se utiliza en múltiples disciplinas para analizar patrones de variabilidad y tomar decisiones fundamentadas. En estudios experimentales, Lee et al. (2015) destacan su uso para representar la precisión y consistencia en datos normales, diferenciando entre desviación estándar (SD) y error estándar de la media (SEM) en análisis inferenciales.

Si bien la desviación estándar es una herramienta poderosa, su sensibilidad a valores atípicos puede afectar su interpretación, especialmente en distribuciones asimétricas. Mahajan (2014)

recomienda combinarla con otras métricas, como el rango intercuartílico, para obtener una visión más completa de la dispersión en datos extremos.

La fórmula matemática para calcular la desviación estándar es:

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

Donde:

x_i = cada valor de la muestra

$\bar{x}$ = media de la muestra

n = tamaño de la muestra

La desviación estándar nos proporciona una medida de dispersión que nos permite comprender cuánto se desvían los valores individuales del conjunto con respecto a su media. Valores más altos de desviación estándar indican una mayor dispersión de los datos.

Ejemplo:

Supongamos que tenemos el siguiente conjunto de datos que representa las edades (en años) de cinco personas: 30, 35, 40, 45, 50.

Solución:

Calculamos la media de estos datos:

$$\bar{x} = \frac{30 + 35 + 40 + 45 + 50}{5}$$

$$\bar{x} = 40$$

Calculamos la varianza utilizando la fórmula:

$$s^2 = \frac{\sum_{i=1}^5 (30-40)^2 + (35-40)^2 + (40-40)^2 + (45-40)^2 + (50-40)^2}{5-1}$$

$$s^2 = \frac{250}{4}$$

$$s^2 = 62.50$$

$$s \approx 7.91$$

Por lo tanto, la desviación estándar de estas edades es aproximadamente 7.91 años. Esto indica que las edades en este conjunto están dispersas alrededor de su media, con una variabilidad de aproximadamente 7.91 años.

Estos ejercicios te ayudarán a practicar el cálculo de la desviación estándar y a comprender cómo esta medida de dispersión nos proporciona información valiosa sobre la variabilidad de los datos en diferentes contextos y situaciones, manteniendo la unidad de medida original de los datos:

- **Datos de Alturas:** Considera el siguiente conjunto de datos que representa las alturas (en centímetros) de una muestra de estudiantes: 160, 165, 170, 175, 180. Calcula la desviación estándar de estas alturas.
- **Datos de Tiempos:** Se registraron los siguientes tiempos (en minutos) que tomó completar una actividad para diferentes personas: 10, 12, 15, 9, 11. Calcula la desviación estándar de estos tiempos.
- **Datos de Temperaturas:** Se registraron las siguientes temperaturas (en grados Celsius) durante una semana: 20, 22, 18, 25, 21. Calcula la desviación estándar de estas temperaturas.

4.5. Coeficiente de variación

El coeficiente de variación (CV) es una medida de dispersión relativa que se utiliza para comparar la variabilidad de conjuntos de datos con diferentes escalas de medida o magnitudes. Su principal utilidad

radica en su capacidad para normalizar la dispersión respecto a la media, lo que lo convierte en una herramienta indispensable para analizar conjuntos de datos con unidades o escalas distintas.

El coeficiente de variación se calcula dividiendo la desviación estándar entre la media y multiplicando el resultado por 100 para expresarlo como un porcentaje. Esto permite determinar qué conjunto de datos presenta una mayor variabilidad relativa en relación con su media, proporcionando un contexto más claro para la comparación.

En el análisis estadístico, el CV se utiliza ampliamente en diversas disciplinas. Según Kesteven (1946), es una herramienta eficaz para medir la variabilidad relativa y es particularmente útil en campos como la biología y la economía, donde las diferencias de escala son comunes. Por otro lado, Lovie (2005) destaca que el CV es independiente de las unidades de medición, lo que facilita comparaciones entre variables heterogéneas.

La fórmula matemática para calcular el coeficiente de variación se expresa como:

$$CV = \frac{S}{\bar{x}} \cdot 100\%$$

El coeficiente de variación se expresa como un porcentaje para facilitar la interpretación y la comparación entre diferentes conjuntos de datos. Cuanto mayor sea el coeficiente de variación, mayor será la variabilidad relativa de los datos en relación con su media.

Ejemplo:

Supongamos que tenemos dos conjuntos de datos que representan el peso de estudiantes en dos grupos diferentes, expresados en kilogramos:

Grupo 1: 50 kg, 55 kg, 60 kg, 65 kg, 70 kg

Grupo 2: 60 kg, 65 kg, 70 kg, 75 kg, 80 kg

Calcularemos el coeficiente de variación para cada grupo para comparar su variabilidad relativa.

Para el Grupo 1:

Calculamos la media:

$$\bar{x} = \frac{50 + 55 + 60 + 65 + 70}{5}$$
$$\bar{x} = 60$$

Calculamos la varianza utilizando la fórmula:

$$s^2 = \frac{\sum_{i=1}^5 (50 - 60)^2 + (55 - 60)^2 + (60 - 60)^2 + (65 - 60)^2 + (70 - 60)^2}{5 - 1}$$

$$s^2 = \frac{250}{4}$$

$$s^2 = 62.50$$

Calculamos la desviación estándar:

$$s \approx 7.91$$

Calculamos el coeficiente de variación:

$$CV = \frac{7.91}{62.50} \cdot 100\%$$

$$CV \approx 12.66\%$$

Para el Grupo 2:

Calculamos la media:

$$\bar{x} = \frac{60 + 65 + 70 + 75 + 80}{5}$$
$$\bar{x} = 70$$

Calculamos la varianza utilizando la fórmula:

$$s^2 = \frac{\sum_{i=1}^5 (60 - 70)^2 + (65 - 70)^2 + (70 - 70)^2 + (75 - 70)^2 + (80 - 70)^2}{5 - 1}$$

$$s^2 = \frac{250}{4}$$

$$s^2 = 62.50$$

Calculamos la desviación estándar:

$$s \approx 7.91$$

Calculamos el coeficiente de variación:

$$CV = \frac{7.91}{70} \cdot 100\%$$

$$CV \approx 11.29\%$$

En este ejemplo, podemos observar que el coeficiente de variación para el Grupo 1 es del 12.66%, mientras que para el Grupo 2 es del 11.29%. Esto indica que el Grupo 1 tiene una mayor variabilidad relativa en el peso de los estudiantes en comparación con su media, en comparación con el Grupo 2.

Estos ejercicios te ayudarán a practicar el cálculo del coeficiente de variación y a comprender cómo esta medida de dispersión relativa es útil para comparar la variabilidad entre diferentes conjuntos de datos con distintas escalas o unidades de medida:

- **Datos de Velocidades:** Se registraron las siguientes velocidades (en km/h) de varios vehículos en dos grupos:
Grupo A: 80, 85, 90, 95, 100
Grupo B: 60, 70, 80, 90, 100
Calcula el coeficiente de variación para cada grupo y compara la variabilidad relativa en las velocidades.
- **Datos de Precios:** Considera los precios de varios productos en dos tiendas diferentes durante una semana:
Tienda 1: \$10, \$12, \$15, \$20, \$25

Tienda 2: \$8, \$10, \$12, \$18, \$22

Calcula el coeficiente de variación para cada tienda y analiza la variabilidad relativa en los precios.

- **Datos de Tiempos:** Se registraron los siguientes tiempos (en minutos) que tomó completar una actividad para diferentes personas en dos grupos:

Grupo X: 15, 20, 25, 30, 35

Grupo Y: 10, 15, 20, 25, 30

Calcula el coeficiente de variación para cada grupo y evalúa la variabilidad relativa en los tiempos.

4.6. Caso de Estudio Simulado: Análisis de Costos en Proyecto de Desarrollo de Viviendas

Una empresa de construcción está desarrollando un nuevo barrio residencial que incluye diferentes tipos de viviendas. Para optimizar la asignación de recursos y establecer estrategias de precios, la empresa necesita analizar la variabilidad en los costos de construcción.

Se recopilan los costos totales de construcción (en millones de dólares) de varios proyectos anteriores de cada tipo de vivienda. Se muestran en la Tabla 28.

Tabla 28

Tipos de vivienda y los costos de construcción

Tipo de Vivienda	Costos de Construcción (Millones de Dólares)
Apartamentos	1.2, 1.3, 1.4, 1.2, 1.3, 1.1, 1.4, 1.2, 1.3, 1.2, 1.3
Casas Adosadas	1.8, 1.7, 2.0, 1.9, 1.85, 1.75, 1.95, 1.8, 1.9, 1.85
Chalets	2.5, 2.7, 2.8, 3.0, 2.6, 2.9, 2.55, 2.65, 2.75, 2.85

Se calculan las desviaciones de cada tipo de vivienda como se muestran las Tablas 29, 30 y 31.

Tabla 29*Análisis de Medidas de Dispersión de los Apartamentos*

APARTAMENTOS			
N	Xi	Xi-Xmedia	(Xi-Xmedia)²
1	1,2	-0,06	0
2	1,3	0,04	0
3	1,4	0,14	0,02
4	1,2	-0,06	0
5	1,3	0,04	0
6	1,1	-0,16	0,03
7	1,4	0,14	0,02
8	1,2	-0,06	0
9	1,3	0,04	0
10	1,2	-0,06	0
11	1,3	0,04	0
	13,9		0,0856

Tabla 30*Análisis de Medidas de Dispersión de las casas adosadas*

CASAS ADOSADAS			
N	Xi	Xi-Xmedia	(Xi-Xmedia)²
1	1,8	-0,05	0,0025
2	1,7	-0,15	0,0225
3	2	0,15	0,0225
4	1,9	0,05	0,0025
5	1,85	0	0
6	1,75	-0,1	0,01
7	1,95	0,1	0,01
8	1,8	-0,05	0,0025
9	1,9	0,05	0,0025
10	1,85	0	0
	18,5		0,075

Tabla 31*Análisis de Medidas de Dispersión de los Chalets*

CHALETS			
N	Xi	Xi-Xmedia	(Xi-Xmedia)²
1	2,5	-0,23	0,0529
2	2,7	-0,03	0,0009
3	2,8	0,07	0,0049
4	3	0,27	0,0729
5	2,6	-0,13	0,0169
6	2,9	0,17	0,0289
7	2,55	-0,18	0,0324
8	2,65	-0,08	0,0064
9	2,75	0,02	0,0004
10	2,85	0,12	0,0144
	27,3		0,231

1. Resumen de la Media y la Desviación Estándar (unidad en millones USD \$):

- Apartamentos: Media = 1.26, Desviación Estándar = 0.09
- Casas Adosadas: Media = 1.85, Desviación Estándar = 0.09
- Chalets: Media = 2.73, Desviación Estándar = 0.16

2. Resumen del Coeficiente de Variación (CV):

- Apartamentos: $CV = (0.09 / 1.26) \times 100\% \approx 7\%$
- Casas Adosadas: $CV = (0.09 / 1.85) \times 100\% \approx 5\%$
- Chalets: $CV = (0.16 / 2.73) \times 100\% \approx 6\%$

Resultados:

1. Los apartamentos muestran una variabilidad moderada en los costos de construcción, con un CV de aproximadamente 7%.
2. Las casas adosadas y los chalets presentan una menor variabilidad en los costos (CV alrededor del 5 y 6%), lo que indica una mayor previsibilidad en los gastos de construcción.

Conclusiones y Decisiones:

- La empresa podría considerar revisar y optimizar los procesos de construcción de apartamentos para reducir la variabilidad en los costos.
- Para las casas adosadas y los chalets, la empresa puede continuar con su enfoque actual, dado que la variabilidad en los costos es relativamente baja.
- Estos análisis ayudan a la empresa a establecer estrategias de precios más precisas y a mejorar la gestión de riesgos y recursos.

Este caso simulado demuestra cómo las medidas de dispersión, especialmente la desviación estándar y el coeficiente de variación, pueden ser utilizadas para analizar y tomar decisiones informadas en proyectos de construcción.

Ejercicios a desarrollar del Capítulo 4

1. Calcula el rango de las alturas de estos edificios. Alturas (en metros) de cinco edificios: 120, 150, 110, 130, 140.
2. Determina el rango de duración de estos proyectos. Duración (en meses) de proyectos: 12, 18, 15, 20, 14.
3. Encuentra el rango de los costos de construcción. Costos (en miles de dólares) de diferentes proyectos: 500, 600, 550, 650, 700.
4. Calcula el rango de las temperaturas. Temperaturas máximas diarias (en °C) en una semana: 32, 35, 31, 33, 34, 30, 36.
5. Determina el rango de la cantidad de concreto utilizado. Cantidad de concreto (en m³) utilizado en diferentes días: 20, 15, 25, 18, 22.
6. Encuentra el rango de los anchos de estos tramos. Anchos (en metros) de diferentes tramos de carretera: 3.5, 4.0, 3.8, 4.2, 3.6.
7. Calcula el rango de las alturas de los puentes. Alturas (en metros) de varios puentes: 50, 55, 60, 45, 65.
8. Determina el rango de profundidad de las fundiciones: Profundidades (en metros) de las fundiciones de edificios: 2, 2.5, 3, 1.8, 2.2.
9. Encuentra el rango de las áreas de las habitaciones: Áreas (en m²) de habitaciones en un diseño: 30, 35, 28, 32, 34.
10. Calcula el rango de los caudales de los ríos: Caudales (en m³/s) de diferentes ríos: 80, 120, 100, 90, 110.
11. Se te proporcionan los costos (en miles de dólares) de un material específico de una muestra de diferentes proveedores. Los costos son los siguientes: 10, 12, 15, 11, 14. Calcular la varianza de los costos de este material para evaluar la dispersión de los precios entre los diferentes proveedores e interprete los resultados.
12. Se te proporcionan los tiempos de duración (en meses) de cinco proyectos de construcción distintos, que son los siguientes: 6, 8, 7, 9, 10. Calcular la varianza de los tiempos de construcción

de estos proyectos para entender mejor cuánto varían en duración.

- 13.** Eres un ingeniero civil encargado de evaluar la calidad del concreto suministrado por diferentes proveedores para un importante proyecto de construcción. Se ha realizado una serie de pruebas de resistencia a la compresión en muestras de concreto de cinco proveedores diferentes. Los resultados de las pruebas, expresados en megapascuales (MPa), se muestran en la Tabla 32.

Tabla 32

Resistencia a la compresión en muestras de concreto

Proveedor	Frecuencia (fa)
A	30
B	35
C	32
D	34
E	33

- a) Calcula la media de la resistencia a la compresión del concreto.
 - b) Determina la varianza de la resistencia a la compresión para evaluar la consistencia entre los diferentes proveedores.
 - c) Basado en los resultados, discute la confiabilidad de los proveedores en términos de la calidad del concreto.
- 14.** Calcula el coeficiente de variación de los costos de los materiales: Costos (en miles de dólares) de un material en diferentes proveedores: 10, 12, 15, 11, 14.
- 15.** Determina el coeficiente de variación de las alturas de estos edificios: Alturas (en metros) de edificios: 120, 150, 130, 140, 135.
- 16.** Calcula el coeficiente de variación de la duración de estos proyectos: Duración (en meses) de proyectos: 12, 14, 16, 18, 15.

- 17.** Determina el coeficiente de variación de la cantidad de concreto utilizado: Cantidad de concreto (en m^3) utilizado en diferentes días: 20, 22, 25, 18, 21.
- 18.** Calcula el coeficiente de variación de las temperaturas: Temperaturas máximas diarias (en $^{\circ}\text{C}$) en una semana: 32, 35, 31, 33, 34.
- 19.** Determina el coeficiente de variación de los anchos de estos tramos: Anchos (en metros) de diferentes tramos de carretera: 3.5, 4.0, 3.8, 4.2, 3.6.
- 20.** Calcula el coeficiente de variación de las profundidades de las fundaciones Profundidades (en metros) de las fundaciones de edificios: 2, 2.5, 3, 1.8, 2.2.
- 21.** Determina el coeficiente de variación de las áreas de las habitaciones: Áreas (en m^2) de habitaciones en un diseño: 30, 35, 28, 32, 34.
- 22.** Calcula el coeficiente de variación de los caudales de los ríos: Caudales (en m^3/s) de diferentes ríos: 80, 120, 100, 90, 110.
- 23.** Determina el coeficiente de variación de las poblaciones en estas urbanizaciones. Poblaciones de cinco urbanizaciones: 5000, 7000, 6000, 8000, 6500.

Conclusiones del Capítulo 4

El capítulo cuatro permitió comprender el papel fundamental de las medidas de dispersión en el análisis estadístico, ya que proporcionaron información clave sobre la variabilidad y la consistencia de los datos. Mientras la media aritmética indicó el valor central de un conjunto de datos, las medidas de dispersión revelaron el grado en que los datos se alejaron de ese valor, permitiendo una interpretación más precisa y detallada en diversos contextos. Sin un análisis adecuado de la dispersión, cualquier evaluación estadística quedaría incompleta, lo que podría conducir a conclusiones erróneas o decisiones mal fundamentadas.

En la ingeniería civil y la arquitectura, la aplicación de estas medidas resultó clave para la gestión eficiente de costos, tiempos y calidad de los materiales. Se evidenció cómo la variabilidad en los costos impactó directamente en el presupuesto general de un proyecto, afectando su planificación y ejecución. Al analizar la dispersión en los precios de los materiales, se hizo evidente la necesidad de contar con estrategias para minimizar los efectos de fluctuaciones imprevistas en los costos.

Asimismo, comprender la dispersión en los tiempos de construcción permitió mejorar la asignación de recursos y optimizar la programación de actividades. La variabilidad en la duración de las distintas fases de un proyecto puede generar retrasos y costos adicionales, por lo que las herramientas estadísticas estudiadas en este capítulo brindaron un marco para prever estos escenarios y mitigarlos mediante una planificación más precisa.

En cuanto a la resistencia de los materiales, se demostró que la variabilidad en sus propiedades influyó en la seguridad y durabilidad de las estructuras, lo que destacó la importancia de controlar estos factores en el diseño y la construcción. La desviación estándar y el coeficiente de variación resultaron herramientas esenciales para evaluar si las diferencias en la resistencia de un material como el concreto se mantenían dentro de los límites aceptables o si representaban un riesgo potencial para la integridad estructural.

El estudio de caso aplicado reafirmó la utilidad de estas herramientas estadísticas en escenarios reales, mostrando cómo el análisis de costos en un proyecto de desarrollo de viviendas ayudó a identificar riesgos y a realizar ajustes estratégicos en la planificación. La capacidad de evaluar la variabilidad de los datos permitió tomar decisiones informadas y optimizar la gestión de los recursos. En la toma de decisiones, no solo es importante conocer un valor promedio, sino también comprender qué tan dispersos están los datos en torno a ese valor para hacer proyecciones confiables y reducir el margen de error en el cálculo de recursos, tiempos y costos.

El capítulo también destacó la interrelación entre las medidas de dispersión y otras herramientas estadísticas. Se demostró que la combinación de análisis gráfico con medidas de dispersión puede brindar una comprensión más completa de los datos, facilitando la detección de valores atípicos y la identificación de patrones que no serían evidentes con un solo tipo de análisis.

En definitiva, el capítulo proporcionó una visión integral sobre el uso de medidas como el rango, la varianza, la desviación estándar y el coeficiente de variación, enfatizando la necesidad de seleccionar la más adecuada según el contexto y los objetivos del análisis. Se concluyó que la correcta interpretación de la dispersión de los datos es un elemento esencial en la toma de decisiones dentro de la ingeniería civil y la arquitectura, contribuyendo a una planificación más precisa y a la ejecución eficiente de proyectos. Además, el manejo adecuado de estas herramientas permitirá a los profesionales del área mejorar la estabilidad financiera de sus proyectos, anticipar posibles desviaciones y garantizar estructuras más seguras y eficientes.

Capítulo 5

REPRESENTACIÓN GRÁFICA DE DATOS

En el análisis de datos, la representación gráfica desempeña un papel preponderante al ofrecer una visualización efectiva y comprensible de la información. Este capítulo se enfoca en explorar las diferentes técnicas y herramientas de representación gráfica de datos, las cuales son fundamentales en el campo de la ingeniería civil y en numerosas disciplinas científicas y comerciales.



5.1. Introducción

Las gráficas permiten comunicar de manera clara y concisa patrones, tendencias y relaciones en los datos, lo que facilita la interpretación y toma de decisiones informadas. Desde diagramas de barras hasta diagramas de dispersión y gráficos de líneas, cada tipo de gráfico tiene su utilidad específica según el tipo de datos y la información que se desea comunicar.

En este capítulo, exploraremos cómo utilizar diferentes tipos de gráficos para representar datos cuantitativos y cualitativos en el contexto de la ingeniería civil. Analizaremos la importancia de elegir el tipo de gráfico adecuado para resaltar la información relevante y ofrecer una perspectiva clara sobre los resultados obtenidos en análisis estadísticos y estudios de investigación.

Además, veremos cómo la representación gráfica complementa las medidas estadísticas discutidas anteriormente, como la media, mediana, moda, varianza y desviación estándar. La combinación de estas herramientas estadísticas con gráficos efectivos nos brinda una comprensión más completa y profunda de los datos, lo que resulta óptimo en la toma de decisiones en el ámbito de la ingeniería civil. A lo largo de este capítulo, exploraremos ejemplos prácticos de cómo crear y analizar diferentes tipos de gráficos, así como su aplicación en la visualización de datos relacionados con proyectos de construcción, gestión de recursos y análisis de riesgos. Al dominar las técnicas de representación gráfica, los profesionales de la ingeniería civil podrán comunicar de manera efectiva resultados complejos y tomar decisiones fundamentadas basadas en datos sólidos y comprensibles.

5.2. Diagramas de barras simples

Un gráfico de barras es una representación visual ampliamente utilizada para datos categóricos, en la que las categorías se representan mediante barras rectangulares cuya longitud o altura es proporcional al valor que representan. Este tipo de gráfico es esencial en estadística para comparar magnitudes de diferentes categorías de manera clara y efectiva. Según Whitaker y Jacobbe (2017), los

gráficos de barras son herramientas fundamentales para comunicar información estadística en contextos educativos y prácticos, especialmente en el análisis de datos categóricos y temporales.

5.2.1. Características

- En el eje Horizontal (X) se representa las categorías, en el eje Vertical (Y) viene representados por los valores o frecuencias de cada categoría.
- El sentido de las barras puede ser verticales u horizontales.
- La altura (en gráficos de barras verticales) o la longitud (en gráficos de barras horizontales) de cada barra es proporcional al valor que representa.
- Las barras están separadas por espacios uniformes para distinguir claramente entre las diferentes categorías.
- Cada barra está etiquetada con la categoría que representa.
- Los valores o frecuencias pueden estar etiquetados en la parte superior de las barras para mayor claridad.
- Este gráfico es ideal para comparar diferentes categorías entre sí, como la comparación de ventas de productos en diferentes meses.
- Muestra cómo se distribuyen los datos entre diferentes categorías, facilitando la identificación de patrones y tendencias.
- Utilizado para mostrar la frecuencia de ocurrencias en cada categoría, como el número de estudiantes que obtuvieron diferentes calificaciones.
- Útil para representar datos categóricos como género, grupos de edad, tipos de productos, etc.

5.2.2. Ventajas

- Facilita la comparación de diferentes categorías de un vistazo.
- Puede utilizarse para una amplia variedad de datos categóricos.
- Fácil de crear y entender, adecuado para audiencias no técnicas.

Ejemplo:

Un arquitecto ha registrado el número de proyectos que ha

completado en los últimos 5 años. Representa estos datos en un gráfico de barras simples.

2019: 4 proyectos

2020: 5 proyectos

2021: 3 proyectos

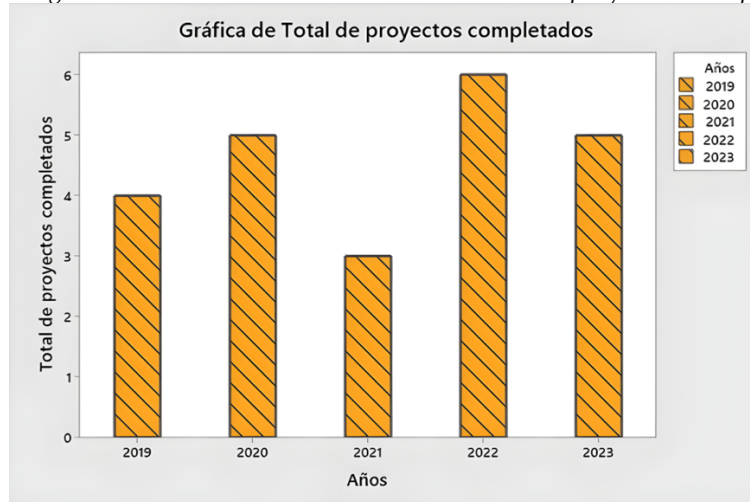
2022: 6 proyectos

2023: 5 proyectos

En la Figura 2, se muestra el eje horizontal, se representan los años (2019 a 2023) y en el eje vertical, el número de proyectos. Cada año tiene una barra cuya altura corresponderá al número de proyectos completados ese año.

Figura 2

Diagrama de barras de la variable de número de proyectos completados



5.3. Diagramas de Barras compuestos

Un gráfico de barras compuesto, también conocido como gráfico de barras apilado, es una variación del gráfico de barras tradicional en el que cada barra se divide en segmentos que representan subcategorías dentro de una categoría principal. Estos gráficos permiten analizar no solo las categorías principales, sino también sus componentes internos, facilitando comparaciones detalladas. Según Heiberger y Robbins (2014), los gráficos de barras apilados son herramientas esenciales en investigaciones donde se requiere comparar la distribución de atributos dentro de categorías múltiples, como en escalas de Likert y otras aplicaciones sociales y

comerciales. Además, Indratmo et al. (2018) destacaron que estos gráficos son efectivos para representar múltiples atributos en tareas de comparación, mejorando la percepción visual en análisis complejos.

5.3.1. Características

- En el eje Horizontal (X) se representa las categorías principales y en el eje Vertical (Y): representa los valores totales de cada categoría principal, divididos en segmentos de subcategorías.
- Cada barra se divide en varios segmentos de diferentes colores o patrones, representando diferentes subcategorías.
- La altura (o longitud, si el gráfico es horizontal) total de la barra representa el valor total de la categoría principal.
- La altura (o longitud) de cada segmento dentro de una barra representa el valor de la subcategoría correspondiente.
- Los valores totales y de subcategorías pueden etiquetarse en la parte superior de las barras o al lado de los segmentos.
- Se incluye una leyenda para explicar el significado de los colores o patrones de los segmentos.

5.3.2. Ventajas

- Facilita la comparación de la estructura interna de diferentes categorías principales.
- Útil para mostrar cómo se distribuyen subcategorías dentro de una categoría principal.
- Permite un análisis más detallado y comprensivo de los datos al mostrar tanto el total de las categorías principales como la contribución de cada subcategoría.
- Ideal para observar cambios en la composición de categorías a lo largo del tiempo.

Ejemplo:

Una empresa constructora ha registrado el número de edificios residenciales y comerciales que ha construido en los últimos 3 años.

Representa estos datos en un gráfico de barras múltiples

2019: Residenciales 4, Comerciales 2

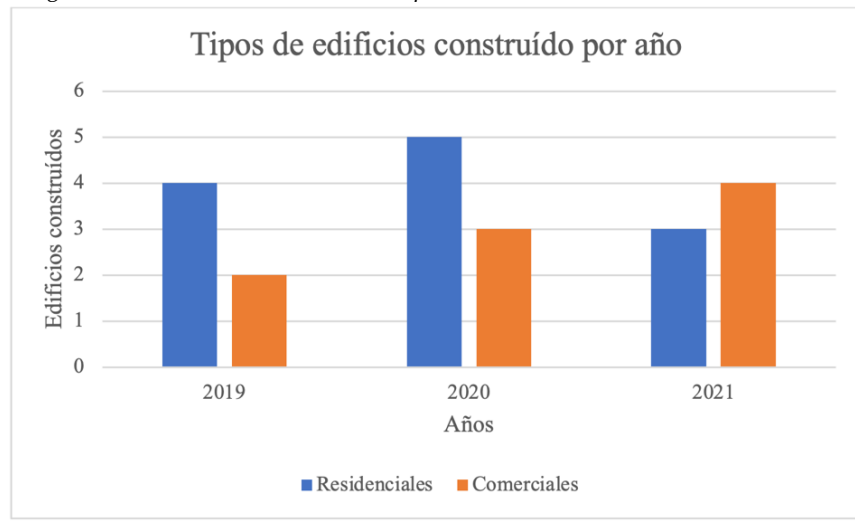
2020: Residenciales 5, Comerciales 3

2021: Residenciales 3, Comerciales 4

En la Figura 3, en el eje horizontal, se representan los años (2019 a 2021) y en el eje vertical, el número de edificios. Para cada año, hay dos barras: una para edificios residenciales y otra para comerciales. La altura de cada barra corresponde al número de edificios construidos de cada tipo en ese año.

Figura 3

Diagrama de barras de la variable tipos de edificios



5.4. Histogramas

Los histogramas son gráficos utilizados en estadística para representar la distribución de frecuencias de un conjunto de datos. Son especialmente adecuados para variables continuas, ya que los datos se dividen en intervalos (clases) y se cuenta el número de observaciones que caen en cada intervalo. A diferencia de los gráficos de barras, que se utilizan para datos categóricos, los histogramas permiten analizar la forma y la distribución de los datos continuos, facilitando la identificación de patrones, tendencias y características importantes. Según Scott (2010), los histogramas son una herramienta clave para visualizar datos y estimar funciones de

densidad, proporcionando una interpretación intuitiva y directa de las distribuciones.

Para construir un histograma, se sigue un proceso que implica los siguientes pasos:

- Se determinan los intervalos o clases en los cuales se agruparán los datos. Estos intervalos deben ser mutuamente excluyentes y abarcar toda la gama de valores de la variable analizada.
- Se cuentan cuántas observaciones caen dentro de cada intervalo. Estas frecuencias se representan en el eje vertical del histograma.
- Se dibujan rectángulos (barras) en el eje horizontal, donde la base de cada barra representa un intervalo de clase y la altura de la barra representa la frecuencia o cantidad de observaciones en ese intervalo.

5.4.1. Características

- El eje Horizontal (X), representa los intervalos de valores (clases) en los que se agrupan los datos y el eje Vertical (Y), representa la frecuencia (el número de veces que ocurren los valores dentro de cada intervalo).
- Las barras en un histograma son adyacentes, sin espacios entre ellas, indicando que los datos son continuos.
- La altura de cada barra representa la frecuencia de los datos dentro de ese intervalo.
- Los intervalos son de igual ancho y cubren todo el rango de los datos.
- La elección del número de intervalos puede afectar la apariencia del histograma y la interpretación de los datos.

5.4.2. Ventajas

- Ideal para visualizar la distribución de datos y observar la forma de la distribución (simétrica, sesgada, etc.).
- Permite identificar patrones como la centralización de datos, dispersión, y presencia de múltiples picos (modas).
- Facilita la identificación de valores atípicos o datos anómalos que

se desvían significativamente del resto de la distribución.

- Se puede utilizar para comparar la distribución de diferentes conjuntos de datos o las mismas variables en diferentes períodos.

Ejemplo 1:

En la Tabla 33 se ha registrado la altura (en metros) de 50 edificios en una ciudad. Representa estos datos en un histograma de frecuencias absolutas.

Tabla 33

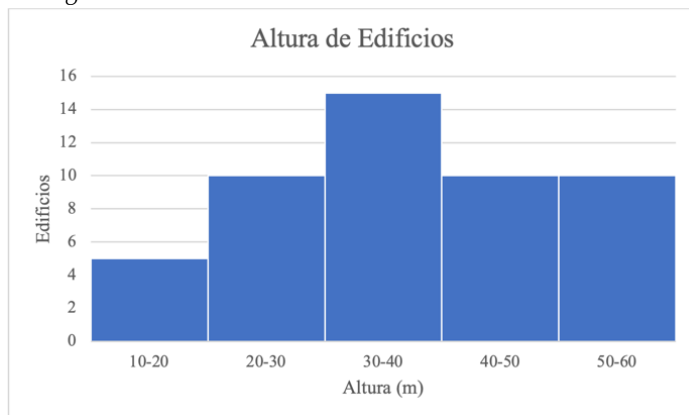
Frecuencias de altura de edificios

Altura (m)	Frecuencia (fa)
10-20	5
20-30	10
30-40	15
40-50	10
50-60	10

En la Figura 4, el histograma muestra cinco barras adyacentes. La altura de cada barra representa la frecuencia de edificios en cada intervalo de altura. Por ejemplo, la barra correspondiente al intervalo 30-40 m sería la más alta, ya que tiene la mayor frecuencia (15 edificios).

Figura 4

Histograma de frecuencias de la altura de los edificios



Ejemplo 2:

En la Tabla 34, se ha registrado la longitud (en metros) de diferentes vigas utilizadas en un proyecto. Representa estos datos en un histograma de frecuencias absolutas.

Tabla 34

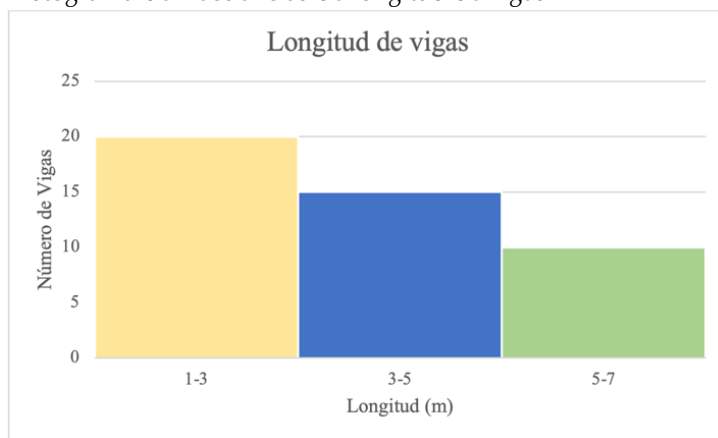
Frecuencias de longitud de vigas

Longitud (m)	Frecuencia (fa)
1-3	20
3-5	15
5-7	10

En la Figura 5, el histograma muestra tres barras adyacentes. La altura de cada barra representa la frecuencia de vigas en cada intervalo de longitud. Por ejemplo, la barra correspondiente al intervalo 1-3 m es la más alta, ya que tiene la mayor frecuencia (20 vigas).

Figura 5

Histograma de frecuencias de longitud de vigas



Al observar el histograma, podemos identificar si la distribución de edades está concentrada en ciertos rangos, si presenta una distribución uniforme o si tiene alguna característica particular, como una cola larga en edades mayores.

5.5. Gráficos Circular, Sector o Pastel

Los gráficos de sectores, también conocidos como gráficos circulares, dividen un círculo en sectores para representar la proporción de cada categoría en relación con el total. Cada sector representa un porcentaje o proporción de la variable total, facilitando la visualización de la distribución relativa de partes dentro de un conjunto de datos. Según Spence y Lewandowsky (1989), los gráficos de sectores son populares debido a su simplicidad y facilidad de comprensión para audiencias generales, aunque tienen limitaciones cuando se trata de comparaciones precisas entre categorías cercanas.

5.5.1. Características

- El círculo está dividido en sectores que representan diferentes categorías de datos.
- Cada sector es proporcional a la frecuencia o porcentaje de la categoría que representa.
- El ángulo de cada sector es proporcional al valor que representa. La suma de todos los ángulos es 360 grados.
- Cada sector generalmente está etiquetado con el nombre de la categoría y su valor o porcentaje correspondiente.
- Puede incluirse una leyenda para describir los colores o patrones usados.
- Los sectores están coloreados o sombreados con diferentes colores o patrones para diferenciarlos fácilmente.

5.5.2. Ventajas

- Muestra la proporción de diferentes categorías dentro de un conjunto de datos total.
- Útil para comparar cómo se distribuye un conjunto de datos entre diferentes categorías.
- Efectivo para visualizar la composición de una variable categórica en términos de porcentaje.
- Comúnmente utilizado en presentaciones e informes para

comunicar la distribución de datos de manera visualmente atractiva.

- Fácil de crear y entender, especialmente adecuado para audiencias no técnicas.

Ejemplo de gráfico circular:

En un proyecto de construcción de un edificio residencial, se ha registrado en la Tabla 35, el presupuesto destinado a diferentes áreas. Representa estos datos en un gráfico de sector.

Tabla 35

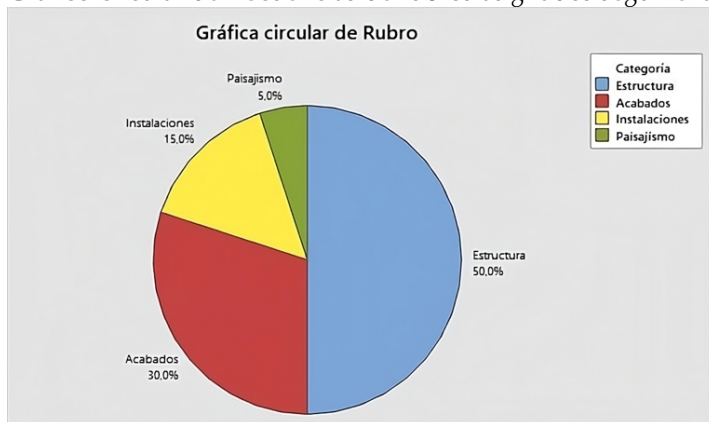
Frecuencias de longitud de vigas

Rubro	Presupuesto
Estructura	\$ 50,000
Acabados	\$ 30,000
Instalaciones	\$ 15,000
Paisajismo	\$ 5,000

En la Figura 6, muestra cuatro sectores circulares. El sector de Estructura es el más grande, ya que representa la mayor parte del presupuesto. Los tamaños de los sectores serían proporcionales a los montos del presupuesto para cada área. Por ejemplo, el sector de Estructura representaría el 50% del total, mientras que el sector de "Paisajismo" sería el más pequeño, representando solo el 5% del total.

Figura 6

Gráfico circular de frecuencias de rubros asignados según el área



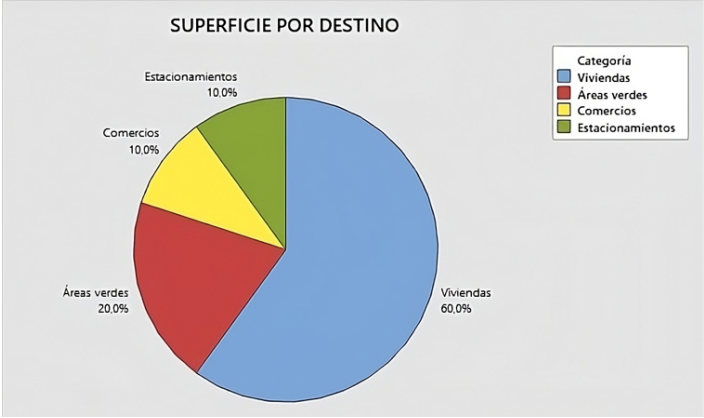
En un proyecto arquitectónico, en la Tabla 36 se ha registrado el área (en metros cuadrados) destinada a diferentes propósitos en un complejo habitacional. Representa estos datos en un gráfico de sector.

Tabla 36
Frecuencias de áreas

Destino	Superficie
Viviendas	6,000 m ²
Áreas verdes	2,000 m ²
Comercios	1,000 m ²
Estacionamientos	1,000 m ²

En la Figura 7, el gráfico de sector muestra cuatro sectores circulares. El sector de Viviendas es el más grande, ya que representa la mayor parte del área total. Los tamaños de los sectores son proporcionales a las áreas destinadas a cada propósito. Por ejemplo, el sector de Viviendas representa el 60% del total, mientras que los sectores de Comercios y Estacionamientos representan cada uno el 10% del total.

Figura 7
Gráfico circular de frecuencias de área



5.6. Polígono de frecuencias acumuladas

El polígono de frecuencias acumuladas, conocido también como ojiva, es una representación gráfica que muestra cómo se acumulan las observaciones en un conjunto de datos. Se distingue entre dos

tipos principales: la ojiva absoluta, que indica el número acumulado de observaciones, y la ojiva relativa, que expresa el porcentaje acumulado de observaciones. Estas representaciones son útiles para entender la distribución general de los datos y detectar patrones acumulativos.

Según Winkler (2021), las ojivas no solo proporcionan una representación visual clara de las frecuencias acumuladas, sino que también ayudan a interpretar la naturaleza de los datos en contextos como estudios socioeconómicos y educativos, donde las distribuciones acumulativas revelan detalles valiosos que otros gráficos no detectan. Además, Cromley y Ye (2006) mencionan que las ojivas son herramientas efectivas para mapear distribuciones estadísticas en contextos espaciales y temporales, especialmente en combinación con otras técnicas visuales como mapas coropléticos.

5.6.1. Características

- El eje Horizontal (X), representa los límites de las clases o intervalos de los datos y el eje Vertical (Y, representa las frecuencias acumuladas absolutas o relativas (porcentajes acumulados).
- Cada punto del gráfico representa el límite superior de una clase y su frecuencia acumulada.
- En una ojiva absoluta, los puntos muestran el número acumulado de observaciones hasta el límite superior de cada intervalo.
- En una ojiva relativa, los puntos muestran el porcentaje acumulado de observaciones hasta el límite superior de cada intervalo.
- Los puntos están conectados por líneas rectas, formando un polígono.
- La línea comienza en el primer límite inferior del primer intervalo con una frecuencia acumulada de 0.

5.6.2. Ventajas

- Facilita la comprensión de la distribución acumulativa de un

conjunto de datos.

- Muestra cómo se acumulan los datos a través de los intervalos.
- Permite comparar de manera efectiva diferentes conjuntos de datos acumulativos.

Ejemplo 1 de Polígono de frecuencias acumuladas:

En la Tabla 37, se ha registrado la altura (en metros) de 50 edificios en una ciudad. Representa estos datos en un polígono de frecuencias acumuladas.

Tabla 37.

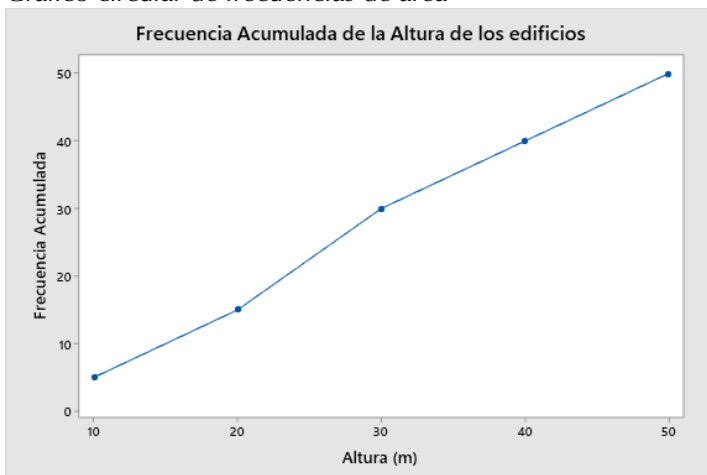
Frecuencias de altura de los edificios

Altura (m)	Frecuencia	Frecuencia Acumulada
10-20	5,00	5,00
20-30	10,00	15,00
30-40	15,00	30,00
40-50	10,00	40,00

En la Figura 8, el gráfico lineal muestra puntos en las alturas 20, 30, 40, 50 y 60 m, con las frecuencias acumuladas correspondientes. Luego, se conectan estos puntos con líneas rectas, formando el polígono de frecuencias acumuladas.

Figura 8

Gráfico circular de frecuencias de área



Ejemplo 2 de Polígono de frecuencias acumuladas:

En la Tabla 38, se ha registrado el peso (en toneladas) de diferentes estructuras en un proyecto. Representa estos datos en un polígono de frecuencias acumuladas.

Tabla 38.

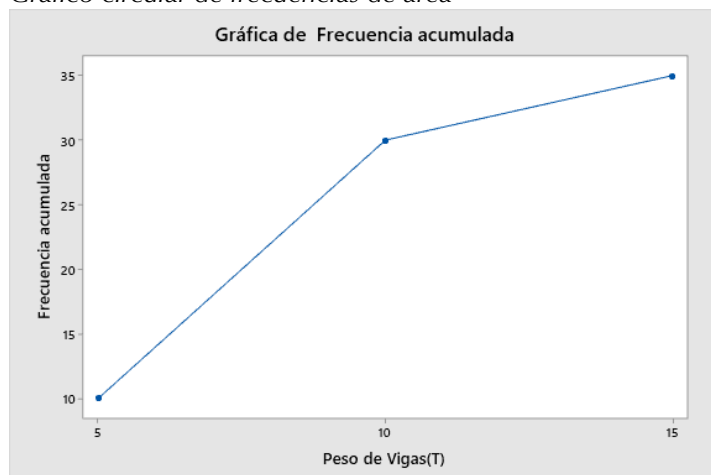
Frecuencias de peso de estructuras

Peso (ton)	Frecuencia	Frecuencia Acumulada
1-5	10	10
5-10	20	30
10-15	5	35

En la Figura 9, el gráfico lineal muestra puntos en los pesos 5, 10 y 15 toneladas, con las frecuencias acumuladas correspondientes. Luego, se conectan estos puntos con líneas rectas, formando el polígono de frecuencias acumuladas.

Figura 9

Gráfico circular de frecuencias de área



Conclusiones del Capítulo 5

La representación gráfica resultó ser una herramienta fundamental para la visualización, interpretación y comunicación de datos en la ingeniería civil. A lo largo del capítulo, se evidenció cómo los distintos tipos de gráficos facilitaron la identificación de patrones, tendencias y relaciones dentro de los datos, complementando las medidas estadísticas y proporcionando un análisis más completo.

El dominio de estas técnicas optimizó la toma de decisiones y mejoró la claridad en la presentación de resultados en contextos técnicos y de gestión. Asimismo, se confirmó el papel esencial de las representaciones gráficas en la práctica profesional, donde una adecuada comunicación de los datos contribuyó a una planificación más eficiente y a la ejecución efectiva de proyectos.

CONCLUSIONES

El libro *Estadística descriptiva aplicada en la Arquitectura e Ingeniería Civil* ha sido concebido como una herramienta fundamental para la formación de futuros profesionales, brindando una base sólida en el uso de la estadística para la toma de decisiones en el ámbito de la construcción y la gestión de proyectos. A lo largo de sus cinco capítulos, se ha destacado la importancia de la estadística como una disciplina esencial para optimizar recursos, mejorar la eficiencia y garantizar la calidad en cada etapa del proceso constructivo.

El desarrollo de los temas ha permitido consolidar un enfoque estructurado, donde la recopilación y organización de datos se presentan como los primeros pasos para cualquier análisis riguroso. Se ha enfatizado que la adecuada selección de métodos de recolección de datos y su correcta representación gráfica no solo facilitan la interpretación de la información, sino que también mejoran la precisión de las decisiones tomadas en el campo de la ingeniería civil y la arquitectura.

Asimismo, se han abordado las medidas de tendencia central y dispersión como herramientas clave para evaluar la variabilidad de los datos y su impacto en la planificación y ejecución de proyectos. Se ha demostrado que el análisis estadístico no solo ayuda a comprender el comportamiento de los datos, sino que también permite identificar patrones y tendencias que pueden influir en la eficiencia y sostenibilidad de las construcciones.

El contenido de este libro trasciende la teoría y ofrece aplicaciones prácticas que buscan preparar a los estudiantes para enfrentar los desafíos actuales de la ingeniería civil. Más allá de los cálculos y métodos estadísticos, se ha resaltado la necesidad de un pensamiento crítico y analítico que permita transformar los datos en información útil y aplicable en la resolución de problemas reales.

Con esta obra, se espera que los lectores adquieran un conocimiento profundo y estructurado sobre la estadística descriptiva y su

aplicación en la arquitectura y la ingeniería civil. La correcta interpretación de los datos es una habilidad fundamental en un sector donde la precisión y la toma de decisiones basadas en evidencia son determinantes para el éxito de cualquier proyecto.

REFERENCIAS

- Albers, M. (2017). Quantitative Data Analysis—In the Graduate Curriculum. *Journal of Technical Writing and Communication*, 47(2), 215–233. <https://doi.org/10.1177/0047281617692067>
- Altman, D., & Bland, J. (2005). Standard deviations and standard errors. *BMJ: British Medical Journal*, 331(7521), 903. <https://doi.org/10.1136/bmj.331.7521.903>
- Bai, Y. (2023). Research on civil engineering cost prediction based on decision tree algorithm. *Academic Journal of Architecture and Geotechnical Engineering*, 5(1), 39-44. <https://doi.org/10.25236/ajage.2023.050107>
- Banerjee, A., Chitnis, U., Jadhav, S., Bhawalkar, J., & Chaudhury, S. (2009). Hypothesis testing, type I and type II errors. *Industrial Psychiatry Journal*, 18(2), 127–131. <https://doi.org/10.4103/0972-6748.62274>
- Cobb, B. R. (2024, July). Attribute statistical process control under nonconstant process deterioration. *Quality and Reliability Engineering International*, 40(5), 2638-2657. <https://doi.org/10.1002/qre.3536>
- Cramer, D., & Howitt, D. (2004). Hypothesis testing. In *The SAGE Dictionary of Statistics* (Vol. 0, pp. 77-77). SAGE Publications, Ltd, <https://doi.org/10.4135/9780857020123>
- Cromley, R., & Ye, Y. (2006). Ogive-based Legends for Choropleth Mapping. *Cartography and Geographic Information Science*, 33(3), 257-268. <https://doi.org/10.1559/152304006779500650>
- Doherty, J., & Hunt, R. J. (2009, March). Two statistics for evaluating parameter identifiability and error reduction. *Journal of Hydrology*, 366(1-4), 119-127. <https://doi.org/10.1016/j.jhydrol.2008.12.018>

- Doncaster, C. P., & Ezard, T. (2017). Key steps to avoiding artistry with significance tests. *PeerJ Preprints*. <https://doi.org/10.7287/peerj.preprints.3394v1>
- Efstathiou, C. E. (2006, July). Estimation of type I error probability from experimental Dixon's "Q" parameter on testing for outliers within small size data sets. *Talanta*, 69(5), 1068-1071. <https://doi.org/10.1016/j.talanta.2005.12.031>
- Fife, D. A. (2019). The Eight Steps of Data Analysis: A Graphical Framework to Promote Sound Statistical Analysis. *Perspectives on Psychological Science*, 15(4), 1054-1075. <https://doi.org/10.1177/1745691620917333>
- Fissan, H., Ristig, S., Kaminski, H., Asbach, C., & Eppele, M. (2014). Comparison of different characterization methods for nanoparticle dispersions before and after aerosolization. *Analytical Methods*, 6(7324), 7324–7334. <https://doi.org/10.1039/c4ay01203h>
- Florian, M., Chapleau, R., Nguyen, S., Achim, C., James-Lefebvre, L., Galarneau, S., Lefebvre, J., & Fisk, C. (1979). Validation and application of an equilibrium-based two-mode urban transportation planning method (EMME). *Transportation Research Record* (pp. 14-23). <https://tinyurl.com/3sy8wu6c>
- Gouveia-Oliveira A. (1996). Estatística e decisão médica [Statistics and medical decision making]. *Acta medica portuguesa*, 9(10-12), 391–396. <https://pubmed.ncbi.nlm.nih.gov/9254540/>
- Hahn, G. J., Doganaksoy, N., & Meeker, W. (2019, August). Statistical intervals, not statistical significance. *Significance*, 16(4), 20-22. <https://doi.org/10.1111/j.1740-9713.2019.01298.x>
- Hall, A. (2011). Experimental Design: Design Experimentation. *Design Issues*, 27(1), 17–26. https://doi.org/10.1162/DESI_a_00074-Hall
- Harrison, S., & Tamaschke, R. (1993). *Statistics for business, economics and management*. Prentice Hall.
- Harrisson, S. (2018). The downside of dispersity: why the standard deviation is a better measure of dispersion in precision

- polymerization. *Polymer Chemistry*, 9(1366–1370).
<https://doi.org/10.1039/C8PY00138C>
- Hayat, M. (2010, May). Understanding Statistical Significance. *Nursing Research*, 59(3), 219–223.
<https://doi.org/10.1097/NNR.0b013e3181dbb2cc>
- Hayslett, H. T., & Murphy, P. (1981). *Measures of Variation. Statistics*. W. H. Allen.
- Heiberger, R., & Robbins, N. B. (2014). Design of diverging stacked bar charts for Likert scales and other applications. *Journal of Statistical Software*, 57(5), 1–32.
<https://doi.org/10.18637/JSS.V057.I05>
- Indratmo, Howorko, L., Boedianto, J. M., & Daniel, B. (2018, September). The efficacy of stacked bar charts in supporting single-attribute and overall-attribute comparisons. *Visual Informatics*, 2(3), 155–165.
<https://doi.org/10.1016/j.visinf.2018.09.002>
- Jakobsen, J., Gluud, C., Winkel, P., Lange, T., & Wetterslev, J. (2014). The thresholds for statistical and clinical significance – a five-step procedure for evaluation of intervention effects in randomised clinical trials. *BMC Medical Research Methodology*, 14(Article 34), 1–12.
<https://doi.org/10.1186/1471-2288-14-34>
- Jankowski, K., & Flannelly, K. (2015). Measures of Central Tendency in Chaplaincy, Health Care, and Related Research. *Journal of Health Care Chaplaincy*, 21(1), 39–49.
<https://doi.org/10.1080/08854726.2014.989799>
- Jiménez-Gamero, I., & Analla, M. (2023). The Importance of Type II Error in Hypothesis Testing. *International Journal of Statistics and Probability*. <https://doi.org/10.5539/ijsp.v12n2p42>
- Kamin, L. F. (2010). Using a Five-Step Procedure for Inferential Statistical Analyses. *American Biology Teacher*, 72(3), 186–188. <https://doi.org/10.1525/abt.2010.72.3.11>
- Kesteven, G. L. (1946). The Coefficient of Variation. *Nature*, 158, 520–521. <https://doi.org/10.1038/158520C0>
- Lee, D. K., In, J., & Lee, S. (2015). Standard deviation and standard

- error of the mean. *Korean Journal of Anesthesiology*, 68, 220-223. <https://doi.org/10.4097/kjae.2015.68.3.220>
- Levy, J. (2008). Use of the Median in Physics and Astronomy. *arXiv: Data Analysis, Statistics and Probability*. <https://doi.org/10.48550/arXiv.0804.0606>
- Liu, B. (2021). Civil engineering based on big data and BIM technology. *Journal of Physics: Conference Series*, 1881, 042007. <https://doi.org/10.1088/1742-6596/1881/4/042007>
- Lovie, P. (2005). Coefficient of Variation. *Encyclopedia of Biostatistics*. <https://doi.org/10.1002/0470013192.BSA107>
- Mahajan, A. (2014, September-October). Why standard deviation as a measure of dispersion needs a mention in a dataset? *Neurology India*, 62(5), 584. <https://doi.org/10.4103/0028-3886.144523>
- Manikandan, S. (2011). Measures of dispersion. *Journal of Pharmacology & Pharmacotherapeutics*, 2(4), 315–316. <https://doi.org/10.4103/0976-500X.85931>
- Mazhar, S. A. (2021). Methods of Data Collection: A Fundamental Tool of Research. *Journal of Integrated Community Health*, 10(1), 6-10. <https://tinyurl.com/2skztrry>
- McCluskey, A., & Lalkhen, A. (2007, June). Statistics I: Data and correlations. *Continuing Education in Anaesthesia, Critical Care & Pain*, 7(3), 95-99. <https://doi.org/10.1093/bjaceaccp/mkm012>
- McGreevy, K., Lipsitz, S., Linder, J., Rimm, E., & Hoel, D. (2009). Using median regression to obtain adjusted estimates of central tendency for skewed laboratory and epidemiologic data. *Clinical Chemistry*, 55(1), 165–169. <https://doi.org/10.1373/clinchem.2008.106260>
- Nguyen, D. H., Nguyen, Q. B., Bui-Tien, T., De Roeck, G., & Wahab, M. A. (2020, October). Damage detection in girder bridges using modal curvatures gapped smoothing method and Convolutional Neural Network: Application to Bo Nghi bridge. *Theoretical and Applied Fracture Mechanics*, 109, 102728. <https://doi.org/10.1016/j.tafmec.2020.102728>

- Parkhurst, D. F. (1998). Arithmetic versus Geometric Means for Environmental Concentration Data. *Environmental Science & Technology*, 32. <https://doi.org/10.1021/ES9834069>
- Rossi, J., & Merenda, P. F. (2010). Variables in Research. *Research Methods in Psychology*, 1-2. <https://doi.org/10.1002/9780470479216.CORPSY1020>
- Sahai, H. (1988). What is a census. *International Journal of Mathematical Education in Science and Technology*, 19(1), 119-124. <https://doi.org/10.1080/0020739880190114>
- Sahu, P.K. (2016). Population and Sample. In: *Applied Statistics for Agriculture, Veterinary, Fishery, Dairy and Allied Fields* (pp. 113-132). Springer.
- Saldanha, L., & Thompson, P. (2002). Conceptions of sample and their relationship to statistical inference. *Educational Studies in Mathematics*, 51(3), 257–270. <https://doi.org/10.1023/A:1023692604014>
- Schadewitz, H., & Blevins, D. R. (1997). A Review of the Data Collection Process. *Journal of Technical Writing and Communication*, 27(4), 417-424. <https://doi.org/10.2190/2GQV-964G-05UQ-V2G6>
- Scott, D. W. (2011, January-February). Histogram. *Wiley Interdisciplinary Reviews: Computational Statistics*, 2(1). <https://doi.org/10.1002/wics.59>
- Seers, K. (2011). Qualitative data analysis. *Evidence-Based Nursing*, 15(2), 2–2. <https://doi.org/10.1136/ebnurs.2011.100352>
- Sergeeva, E. (2020). The Importance of Mathematics for Future Architects and Civil Engineers. *IOP Conference Series: Materials Science and Engineering*, 753. <https://doi.org/10.1088/1757-899X/753/5/052024>
- Sham, P., & Purcell, S. (2014). Statistical power and significance testing in large-scale genetic studies. *Nature Reviews Genetics*, 15, 335–346. <https://doi.org/10.1038/nrg3706>
- Singh, S. (2003). Simple Random Sampling. In *Advanced Sampling Techniques* (71–136). Springer.
- Spence, I., & Lewandowsky, S. (1989). The perception of statistical

- graphs. *Sociological Methods & Research*, 18(2), 200-242.
<https://doi.org/10.1177/0049124189018002002>
- Storey, J. D., & Tibshirani, R. (2003). Statistical significance for genomewide studies. *PNAS*, 100(16), 9440–9445.
<https://doi.org/10.1073/pnas.1530509100>
- Sun, X. (2023). The impact of digital infrastructure on energy-environmental efficiency: empirical evidence from China. *Frontiers in Energy Research*, 11.
<https://doi.org/10.3389/fenrg.2023.1277333>
- Thulasingam, M., & Premarajan, K. (2018). Statistical Analysis: Data Presentation and Statistical Tests. In S. Parija, & V. Kate (eds), *Thesis Writing for Master's and Ph.D. Program* (pp. 113-129). Springer. https://doi.org/10.1007/978-981-13-0890-1_11
- Trafimow, D., Tong, T., Wang, T., Choy, S. T. B., Hu, L., Chen, X., Wang, C., & Wang, Z. (2024). Improving inferential analyses predata and postdata. *Psychological Methods*. Advance online publication. <https://doi.org/10.1037/met0000697>
- Watier, N., Lamontagne, C., & Chartier, S. (2011). What Does the Mean Mean? *Journal of Statistics Education*, 19(2).
<https://doi.org/10.1080/10691898.2011.11889615>
- Wesolowsky, G. O. (1975). Location of the Median Line for Weighted Points. *Environment and Planning A*, 7(2), 163–170. <https://doi.org/10.1068/a070163>
- Whitaker, D., & Jacobbe, T. (2017). Students' Understanding of Bar Graphs and Histograms: Results from the LOCUS Assessments. *Journal of Statistics Education*, 25(2), 90–102.
<https://doi.org/10.1080/10691898.2017.1321974>
- Williams, P. J., & Hooten, M. B. (2016). Combining statistical inference and decisions in ecology. *Ecological Applications*, 26(6), 1930–1942. <https://doi.org/10.1890/15-1593.1>
- Winkler, O. W. (2021). Interpreting the Cumulative Frequency Distribution of Socio-Economic Data. *New Visions in Science and Technology*, Vol. 4.
<https://doi.org/10.9734/bpi/nvst/v4/4236f>
- Xiu, Y. (2024). Application of data mining technology in civil

- engineering systems. *Proceedings of the First International Conference on Science, Engineering and Technology Practices for Sustainable Development*. <https://doi.org/10.4108/eai.17-11-2023.2342713>
- Ye, X., Liu, H., Xiao, X., & Ling, M. (2018). The new concepts of measurement error's regularities and effect characteristics. *Measurement*, 126, October 2018. <https://doi.org/10.1016/j.measurement.2018.05.045>
- Zheng, W. (2011). Space mode and planning of integration of traffic and land use. *Urban Planning Studies*. <https://tinyurl.com/5c4cs2mu>
- Zorić, A. B. (2021). Applied statistics: Basic principles and application. *International Journal of Innovation and Economic Development*, 7(3), 27-33. <https://tinyurl.com/59npc2s8>

RESPUESTAS A LOS EJERCICIOS

Capítulo 1

Sesión 1.1 Identifique en los siguientes enunciados los conceptos de Población

1. **Respuesta:** La población en este estudio son todas las casas construidas en el nuevo desarrollo residencial.

Explicación: La población se refiere al conjunto completo de individuos, objetos o eventos que se quieren estudiar. En este caso, el arquitecto quiere estudiar todas las casas del desarrollo, por lo que esas casas constituyen la población.

2. **Respuesta:** La población en este estudio es el conjunto de todas las construcciones que usaron ese tipo específico de concreto.

Explicación: Aunque el ingeniero solo tomó muestras de 10 construcciones, la población se refiere al conjunto completo de construcciones que usaron ese concreto, no solo a las que se tomaron como muestra.

3. **Respuesta:** La población serían todos los estudiantes de arquitectura de las 5 escuelas de la ciudad.

Explicación: La población incluye a todos los individuos que queremos estudiar. En este caso, queremos la opinión de todos los estudiantes de arquitectura de la ciudad, por lo que ellos constituyen la población.

4. **Respuesta:** La población son todos los puentes peatonales construidos en la ciudad en los últimos 10 años.

Explicación: La población se refiere al conjunto completo de objetos o eventos que queremos estudiar. En este caso, son todos los puentes peatonales construidos en ese período de tiempo.

5. **Respuesta:** La población en este estudio son los 500 clientes de la empresa.

Explicación: La población incluye a todos los individuos u objetos que queremos estudiar. En este caso, la empresa quiere la opinión de todos sus clientes, por lo que esos 500 clientes constituyen la población.

Sesión 1.2 Identifique en los siguientes enunciados los conceptos de Muestra

6. **Respuesta:** La muestra en este estudio son los 15 edificios de oficinas seleccionados por el arquitecto.

Explicación: Una muestra es un subconjunto seleccionado de la población.

En este caso, de los 100 edificios de oficinas (población), el arquitecto seleccionó 15 para su estudio, constituyendo así la muestra.

7. **Respuesta:** La muestra en este estudio son las 20 construcciones seleccionadas por el ingeniero.
Explicación: A pesar de que hay 500 construcciones que usaron el concreto (población), el ingeniero solo seleccionó 20 de ellas para su análisis, lo que constituye la muestra.
8. **Respuesta:** La muestra son los 500 estudiantes seleccionados al azar para responder la encuesta.
Explicación: La muestra es un subconjunto de la población que se elige para representarla. En este caso, de los 10,000 estudiantes (población), se seleccionaron 500 para la encuesta, formando así la muestra.
9. **Respuesta:** La muestra son los 8 edificios residenciales seleccionados para la encuesta.
Explicación: De los 50 edificios residenciales construidos por la empresa (población), se eligieron 8 para realizar la encuesta, lo que constituye la muestra.
10. **Respuesta:** La muestra son las 30 casas seleccionadas para el estudio.
Explicación: Aunque hay 200 casas que usaron la pintura (población), el estudio solo seleccionó 30 de ellas para evaluar la durabilidad, lo que constituye la muestra.

Sesión 1.3 Identifique en los siguientes enunciados los conceptos de Variable

11. **Respuesta:** La variable es cualitativa o categórica.
Explicación: Una variable cualitativa o categórica describe características o categorías y no tiene un orden o medida numérica. En este caso, los tipos de materiales (concreto, acero, madera, ladrillo) son categorías, por lo que la variable es cualitativa.
12. **Respuesta:** La variable es cuantitativa o numérica.
Explicación: Una variable cuantitativa o numérica tiene valores numéricos que representan una cantidad. En este caso, la resistencia a la compresión se mide en MPa, lo que indica una cantidad numérica. Por lo tanto, la variable es cuantitativa.
13. **Respuesta:** La variable es cualitativa o categórica.
Explicación: La clasificación de espacios en una casa (habitaciones, baños, cocina, sala, comedor) representa diferentes categorías o tipos de espacios.

No hay una medida numérica asociada, por lo que la variable es cualitativa.

14. **Respuesta:** La variable es cuantitativa o numérica.
Explicación: La luminosidad se mide en lux, lo que indica una cantidad numérica. Por lo tanto, la variable que representa la luminosidad en diferentes áreas del edificio es cuantitativa.
15. **Respuesta:** La variable es cualitativa o categórica.
Explicación: La clasificación de los tipos de pavimentos (asfalto, concreto, adoquín, grava) representa diferentes categorías o tipos

Sesión 1.4 Identifique en los siguientes enunciados los conceptos de nivel de medición de los datos

16. **Respuesta:** Los materiales representan un nivel de medición nominal, ya que simplemente clasifican los datos en categorías sin un orden específico.
17. **Respuesta:** Los tipos de suelos representan un nivel de medición nominal, ya que se clasifican en categorías sin un orden inherente.
18. **Respuesta:** Los estilos arquitectónicos representan un nivel de medición nominal, ya que son categorías sin un orden específico.
19. **Respuesta:** Las prioridades representan un nivel de medición ordinal, ya que tienen un orden específico, pero la distancia entre las categorías no es uniforme.
20. **Respuesta:** Las categorías de resistencia representan un nivel de medición ordinal, ya que tienen un orden específico de calidad.
21. **Respuesta:** Las áreas representan un nivel de medición ordinal, ya que tienen un orden específico de importancia.
22. **Respuesta:** Las temperaturas representan un nivel de medición de intervalo, ya que tienen un orden y una distancia uniforme entre las mediciones, pero no tienen un verdadero punto de partida.
23. **Respuesta:** Las inclinaciones representan un nivel de medición de intervalo, ya que tienen un orden y una distancia uniforme, pero no un punto de partida absoluto.
24. **Respuesta:** Las intensidades lumínicas representan un nivel de medición de intervalo, ya que tienen un orden y una distancia uniforme entre las mediciones, pero no un punto de partida absoluto.
25. **Respuesta:** Las cantidades de concreto representan un nivel de medición de razón, ya que tienen un orden, una distancia uniforme entre las mediciones y un punto de partida absoluto (0 m³ indica ausencia de concreto).
26. **Respuesta:** Las longitudes representan un nivel de medición de razón, ya que tienen un orden, una distancia uniforme y un punto de partida absoluto (0 m indica ausencia de longitud).
27. **Respuesta:** Los tiempos representan un nivel de medición de razón, ya que

tienen un orden, una distancia uniforme entre las mediciones y un punto de partida absoluto (0 días indica que no ha pasado tiempo).

Capítulo 2

Respuestas a los ejercicios

1. La tabla muestra cuántos vehículos de cada tipo pasaron por el puente en una hora, así como las frecuencias relativas y acumuladas. Por ejemplo, 6 coches pasaron por el puente, lo que representa el 50% de los vehículos registrados.

Tipo de vehículo	Frecuencia Absoluta	Frecuencia Relativa	Frecuencia Acumulada	Frecuencia Relativa Acumulada
Bicicleta	2	0.167 (16.7%)	2	0.167 (16.7%)
Coche	6	0.5 (50%)	8	0.667 (66.7%)
Camión	3	0.25 (25%)	11	0.917 (91.7%)
Moto	1	0.083 (8.3%)	12	1 (100%)
Total	12	1.0 (100%)		

2. La tabla muestra cuántas ventanas de cada tamaño hay en el edificio, así como las frecuencias relativas y acumuladas. Por ejemplo, hay 7 ventanas grandes, lo que representa el 46.7% de todas las ventanas del edificio.

Tamaño de ventana	Frecuencia Absoluta	Frecuencia Relativa	Frecuencia Acumulada	Frecuencia Relativa Acumulada
Pequeña	4	0.267 (26.7%)	4	0.267 (26.7%)
Mediana	4	0.267 (26.7%)	8	0.534 (53.4%)
Grande	7	0.466 (46.6%)	15	1 (100%)
Total	15	1.0 (100%)		

3. El concreto es el material más utilizado en las construcciones recientes, con un 50% de las edificaciones utilizando este material.

Material	Frecuencia Absoluta	Frecuencia Relativa	Frecuencia Acumulada	Frecuencia Relativa Acumulada
Concreto	25	0.50	25	0.50
Ladrillo	10	0.20	35	0.70
Madera	8	0.16	43	0.86
Acero	7	0.14	50	1.00
Total	50	1.0 (100%)		

4. **Rango:** $34 - 30 = 4$.

Número de clases (usando la regla de la raíz cuadrada): $k = \sqrt{8} \approx 3$

Ancho de clase: $4/3 \approx 1.33$ (aproximamos a 1.5).

Resistencia (MPa)	Frecuencia Absoluta	Frecuencia Relativa	Frecuencia Acumulada	Frecuencia Relativa Acumulada
[30 – 32)	4	0.500	4	0.500
[32 – 34)	3	0.375	7	0.875
[34 – 36)	1	0.125	8	1.000
Total	8	1.0 (100%)		

5. **Rango:** $27 - 20 = 7$.

Número de clases (usando la regla de Rice): $k = 2 \times \sqrt{6} \approx 5$

Ancho de clase: $7/5 = 1.4$ (aproximamos a 2).

Altura (m)	Frecuencia Absoluta	Frecuencia Relativa	Frecuencia Acumulada	Frecuencia Relativa Acumulada
20-21	2	0.33	2	0.33
22-23	1	0.17	3	0.50
24-25	2	0.33	5	0.83
26-27	1	0.17	6	1.00
Total	6	1.0 (100%)		

6. Respuesta: 22 edificios tienen una altura superior a 18 metros, sumando las frecuencias absolutas de los rangos 19-20 y 21-22.
7. Respuesta: 30 salones tienen un área entre 62 m^2 y 67 m^2 , sumando las frecuencias absolutas de los rangos 62-63, 64-65 y 66-67.

Capítulo 3

Respuestas a los ejercicios

1. Solución: $\bar{x} = (4.2 + 3.8 + 4.0 + 4.1 + 3.9) / 5 = 20 / 5 = 4$ metros.
2. Solución: $\bar{x} = (12 + 15 + 13 + 14 + 16) / 5 = 70 / 5 = 14$ mil dólares.
3. Solución: $\bar{x} = (50 + 55 + 60 + 65 + 70 + 75) / 6 = 375 / 6 = 62.5$ metros.
4. Solución: $\bar{x} = (12 + 14 + 16 + 18 + 20 + 22 + 24) / 7 = 126 / 7 = 18$ meses.
5. Solución: $\bar{x} = (320 + 340 + 310 + 330 + 350 + 360 + 340 + 320) / 8 = 2770 / 8 = 346.25$ metros cúbicos.
6. Solución: $\bar{x} = (200 + 220 + 210 + 230 + 240 + 250 + 210 + 220 + 230) / 9 = 2010 / 9 = 223.33$ mil dólares.
7. Solución:
 - Media = $\Sigma(xi * fi) / \Sigma fi$
 - Media = $[(2.5 \times 3) + (3.0 \times 5) + (3.5 \times 2)] / (3 + 5 + 2)$
 - Media = $(7.5 + 15 + 7) / 10 = 29.5 / 10$
 - Media = 2.95 m.
8. Solución:
 - Media = $\Sigma(xi * fi) / \Sigma fi$
 - Media = $[(50 \times 4) + (75 \times 6) + (100 \times 2)] / (4 + 6 + 2)$
 - Media = $(200 + 450 + 200) / 12$
 - Media = $850 / 12$
 - Media ≈ 70.83 m².
9. Solución:
 - Media = $\Sigma(xi * fi) / \Sigma fi$
 - Media = $[(200 \times 3) + (250 \times 5) + (300 \times 2)] / (3 + 5 + 2)$
 - Media = $(600 + 1250 + 600) / 10$
 - Media = $2450 / 10$
 - Media = 245 mil \$.
10. Solución:
 - Media = $\Sigma(xi * fi) / \Sigma fi$
 - Media = $[(6 \times 4) + (8 \times 6) + (10 \times 3)] / (4 + 6 + 3)$
 - Media = $(24 + 48 + 30) / 13$
 - Media = $102 / 13$
 - Media ≈ 7.85 meses.
11. Solución:
 - Media = $\Sigma(xi * fi) / \Sigma fi$
 - Media = $[(10 \times 5) + (15 \times 8) + (20 \times 4)] / (5 + 8 + 4)$
 - Media = $(50 + 120 + 80) / 17$
 - Media = $250 / 17$
 - Media ≈ 14.71 toneladas.
12. Solución:
 - Media = $\Sigma(xi * fi) / \Sigma fi$
 - Media = $[(100 \times 10) + (150 \times 15) + (200 \times 5)] / (10 + 15 + 5)$
 - Media = $(1000 + 2250 + 1000) / 30$
 - Media = $4250 / 30$

- Media $\approx 141.67 \text{ m}^3$.
13. La mediana del espesor de los muros es 21.53 cm.
 14. La mediana del área de las fachadas es 205 m².
 15. La mediana de la altura de los pisos es 3.78 m.
 16. La mediana de la cantidad de ventanas es 19.11
 17. La mediana del tiempo de construcción es 20.63 meses
 18. Solución:
 - Ordenar los datos: 120, 130, 135, 140, 150.
 - La mediana es el valor medio: 135 metros.
 19. Proceso:
 - Ordenar los datos: 2.5, 3.0, 3.5, 4.0, 4.5, 5.0.
 - La mediana es el promedio de los dos valores medios: $(3.5 + 4.0) / 2 = 3.75$ hectáreas.
 20. Proceso:
 - Ordenar los datos: 18, 19, 20, 21, 22, 23, 24, 25.
 - La mediana es el promedio de los dos valores medios: $(21 + 22) / 2 = 21.5$ mil dólares.
 21. Proceso:
 - Ordenar los datos: 12, 13, 14, 15, 16, 17, 18.
 - La mediana es el valor medio: 15 meses.
 22. Proceso:
 - Ordenar los datos: 310, 310, 320, 320, 330, 330, 340, 340, 350, 360.
 - La mediana es el promedio de los dos valores medios: $(330 + 330) / 2 = 330$ metros cúbicos.
 23. Proceso:
 - Ordenar los datos: 50, 55, 60, 65, 70, 75, 80, 85, 90.
 - La mediana es el valor medio: 70 metros.
 24. Solución:
 - Ordenar los datos: 100, 120, 130, ?, 150, 160.
 - La mediana es el promedio de los dos valores medios (130 y 150): $(130 + 150) / 2 = 140$ metros.
 25. Solución:
 - Ordenar los datos: <5, 6, 7, ?, >10.
 - La mediana es 7 hectáreas.
 26. Solución:
 - Ordenar los datos: <200, 220, 240, ?, 280, >300.
 - La mediana es el promedio de 240 y 280: $(240 + 280) / 2 = 260$ mil dólares.
 27. Solución:
 - Ordenar los datos: <12, 14, 16, ?, 20, >22.
 - La mediana es 18 meses (promedio de 16 y 20).
 28. Solución:
 - Ordenar los datos: <15, 17, 19, ?, 23, 25, >27.
 - La mediana es 21 toneladas (promedio de 19 y 23).
 29. Solución:
 - Ordenar los datos: <300, 320, 340, ?, 380, 400, >420.
 - La mediana es 360 metros cúbicos (promedio de 340 y 380).

30. Solución:

Tabla de Frecuencias:

Intervalo	Frecuencia	Frecuencia acumulada
100-120m	3	3
120-140m	5	8
140-160m	2	10

Fórmula de Mediana para Datos Agrupados:

$$Mediana = L_{inf} + \left(\frac{\frac{n}{2} - F_{i-1}}{F_i} \right) \times intervalo$$

1. Hallamos $n/2$
 2. Identificamos el valor F que sea igual al valor $n/2$. Caso contrario, los dos F más cercanos.
 3. Identificamos la clase de F_i
 4. Hallamos el límite inferior para F_i y el intervalo
 5. Aplicamos fórmula
1. $N=10$; $N/2 = 5$
 2. $F_i = 8$; $F_{i-1} = 3$
 3. Clase (120 – 140)m
 4. $L_{inf} = 120$; intervalo = 20
 5. Mediana =

$$Mediana = 120 + \left(\frac{\frac{10}{2} - 3}{8} \right) \times 20 = 125 \text{ m}$$

31. Solución:

Tabla de Frecuencias:

Intervalo	Frecuencia	Frecuencia absoluta
1-3 ha	4	4
3-5 ha	6	10
5-7 ha	2	12

$$Mediana = L_{inf} + \left(\frac{\frac{n}{2} - F_{i-1}}{F_i} \right) \times intervalo$$

1. $N=12$; $N/2 = 6$
2. $F_i = 10$; $F_{i-1} = 4$
3. Clase (3-5)Ha

4. $L_{inf} = 3$; intervalo = 2
5. Mediana =

$$Mediana = 3 + \left(\frac{\frac{12}{2} - 4}{10} \right) \times 2 = 3.4 \text{ Ha}$$

32. Solución:

Tabla de Frecuencias:

Intervalo	Frecuencia	Frecuencia absoluta
6-12 meses	3	3
12-18 meses	7	10
18-24 meses	5	15

$$Mediana = L_{inf} + \left(\frac{\frac{n}{2} - F_{i-1}}{F_i} \right) \times intervalo$$

1. $N=15$; $N/2 = 7.5$
2. $F_i = 10$; $F_{i-1} = 3$
3. Clase (12-18) meses
4. $L_{inf} = 12$; intervalo = 6
5. Mediana =

$$Mediana = 12 + \left(\frac{\frac{15}{2} - 3}{10} \right) \times 6 = 14,7 \text{ meses}$$

33. Solución:

Tabla de Frecuencias:

Intervalo	Frecuencia	Frecuencia absoluta
10-20k	4	4
20-30k	8	12
30-40k	3	15

1. $N=15$; $N/2 = 7.5$
2. $F_i = 12$; $F_{i-1} = 4$
3. Clase (20-30) k
4. $L_{inf} = 20$; intervalo = 10
5. Mediana =

$$Mediana = 20 + \left(\frac{\frac{15}{2} - 4}{12} \right) \times 10 = 22,92K$$

34. Solución:

Tabla de Frecuencias:

Intervalo	Frecuencia	Frecuencia absoluta
100-200 m³	5	5
200-300 m³	10	15
300-400 m³	5	20

1. $N=20$; $N/2 = 10$
2. $F_i = 15$; $F_{i-1} = 5$
3. Clase (200-300) m³
4. $L_{inf} = 200$; intervalo = 100
5. Mediana =

$$Mediana = 200 + \left(\frac{\frac{20}{2} - 5}{15} \right) \times 100 = 233,33 m^3$$

35. Solución:

Tabla de Frecuencias:

Intervalo	Frecuencia	Frecuencia absoluta
30-50 m	4	4
50-70 m	8	12
70-90 m	3	15

1. $N=15$; $N/2 = 7.5$
2. $F_i = 12$; $F_{i-1} = 4$
3. Clase (50 - 70) m
4. $L_{inf} = 50$; intervalo = 20
5. Mediana =

$$Mediana = 50 + \left(\frac{\frac{15}{2} - 4}{12} \right) \times 20 = 55,83 m$$

36. Solución:

Tabla de Frecuencias:

Material	Frecuencia
Hormigón	2
Acero	3
Madera	2

Moda: Acero (es el material más frecuente).

37. Solución:

Tabla de Frecuencias:

Estilo	Frecuencia
Moderno	3
Clásico	2
Minimalista	1

Moda: Moderno.

38. Solución:

Tabla de Frecuencias:

Pavimento	Frecuencia
Asfalto	2
Adoquín	3
Concreto	2

Moda: Adoquín.

39. Solución:

Tabla de Frecuencias:

Color	Frecuencia
Rojo	2
Azul	2
Verde	3

Moda: Verde.

40. Solución:

Tabla de Frecuencias:

Iluminación	Frecuencia
LED	3
Fluorescente	2
Incandescente	1

Moda: LED.

41. Solución:

Tabla de Frecuencias:

Estilo	Frecuencia
Panorámico	3
Clásico	2
Abatible	1

Moda: Panorámico.

42. Solución: La moda es 120 metros, ya que es la altura que aparece con mayor frecuencia.
43. Solución: La moda es 200 metros, siendo la longitud más frecuente.
44. Solución: La moda es 2.5 hectáreas, ya que es el área que se repite con mayor frecuencia.
45. Solución: La moda es 80 m³/s, siendo el caudal que aparece más veces.
46. Solución: La moda es 10 y 15, ya que ambos números de pisos aparecen con la misma frecuencia máxima.
47. Solución: La moda es 12 metros, siendo el ancho que se repite más frecuentemente.

48. Solución:

Tabla de Frecuencias:

Intervalo de Altura	Frecuencia
[100-150]	5
[150-200]	8
[200-250]	3

Moda: Intervalo [150-200], ya que tiene la mayor frecuencia.

49. Solución:

Tabla de Frecuencias:

Intervalo de Longitud	Frecuencia
[100-200]	4
[200-300]	7
[300-400]	2

Moda: Intervalo [200-300], siendo el más frecuente.

50. Solución:

Tabla de Frecuencias:

Intervalo de Área	Frecuencia
[1-3]	6
[3-5]	9
[5-7]	4

Moda: Intervalo [3-5], ya que es el más repetido.

51. Solución:

Tabla de Frecuencias:

Intervalo de Caudal	Frecuencia
[50-100]	5
[100-150]	10
[150-200]	3

Moda: Intervalo [100-150], siendo el más común.

52. Solución:

Tabla de Frecuencias:

Intervalo de Pisos	Frecuencia
[0-10]	8
[10-20]	12
[20-30]	5

Moda: Intervalo [10-20], ya que tiene la mayor cantidad de edificios.

53. Solución:

Tabla de Frecuencias:

Intervalo de Ancho	Frecuencia
[10-15]	7
[15-20]	11
[20-25]	4

Moda: Intervalo [15-20], siendo el más frecuente.

Capítulo 4

Respuestas a los ejercicios

1. **Respuesta:** Rango = $150 - 110 = 40$ metros.
2. **Respuesta:** Rango = $20 - 12 = 8$ meses.
3. **Respuesta:** Rango = $700 - 500 = 200$ mil dólares.
4. **Respuesta:** Rango = $36 - 30 = 6$ °C.
5. **Respuesta:** Rango = $25 - 15 = 10$ m³.
6. **Respuesta:** Rango = $4.2 - 3.5 = 0.7$ metros.
7. **Respuesta:** Rango = $65 - 45 = 20$ metros.
8. **Respuesta:** Rango = $3 - 1.8 = 1.2$ metros.
9. **Respuesta:** Rango = $35 - 28 = 7$ m².
10. **Respuesta:** Rango = $120 - 80 = 40$ m³/s.

11. Procedimiento:

Cálculo de la Media:

Calcula la media (promedio) de los costos de los materiales.

Media = (Suma de todos los costos) / (Número total de proveedores)

$$\bar{X} = \frac{10 + 12 + 15 + 11 + 14}{5} = \frac{62}{5} = 12.4 \text{ miles de dólares}$$

Cálculo de la Varianza:

La varianza mide cuánto se desvían los costos individuales de la media.

$$\text{Varianza } (s^2) = \sum \frac{(X_i - \bar{X})^2}{n - 1}$$

En este caso:

X_i : el costo de los materiales en cada proyecto.

$\bar{X}$: la media del costo de los materiales en cada proyecto.

n : la cantidad de proyectos analizados

$$s^2 = \frac{(10 - 12.4)^2 + (12 - 12.4)^2 + (15 - 12.4)^2 + (11 - 12.4)^2 + (14 - 12.4)^2}{5 - 1}$$

$$s^2 = \frac{(-2.4)^2 + (-0.4)^2 + (2.6)^2 + (-1.4)^2 + (1.6)^2}{5 - 1}$$

$$s^2 = \frac{5.76 + 0.16 + 6.76 + 1.96 + 2.56}{4} = \frac{17.2}{4} = 4.3 \text{ miles de dólares}^2$$

Respuesta: La varianza de los costos de los materiales es de 4.3 (miles de dólares)².

Interpretación: Una varianza de 4.3 (miles de dólares)² indica que hay una variabilidad moderada en los precios ofrecidos por los diferentes proveedores. Este análisis ayuda en la toma de decisiones en la selección de proveedores y la gestión presupuestaria en proyectos de construcción y diseño arquitectónico.

12. Procedimiento:

Cálculo de la Media:

Primero, calcula la media (promedio) de los tiempos de construcción.

Media = (Suma de todos los tiempos de construcción) / (Número total de proyectos)

$$\bar{X} = \frac{6 + 8 + 7 + 9 + 10}{5} = \frac{40}{5} = 8 \text{ meses}$$

Cálculo de la Varianza:

La varianza se calcula como el promedio de los cuadrados de las diferencias entre cada tiempo de construcción y la media.

$$\text{Varianza } (s^2) = \sum \frac{(X_i - \bar{X})^2}{n - 1}$$

En este caso:

X_i : el tiempo de construcción de cada proyecto.

$\bar{X}$: la media entre los tiempos de construcción de cada proyecto.

n : la cantidad de proyectos analizados

$$s^2 = \frac{(6 - 8)^2 + (8 - 8)^2 + (7 - 8)^2 + (9 - 8)^2 + (10 - 8)^2}{5 - 1}$$

$$s^2 = \frac{(2)^2 + (0)^2 + (-1)^2 + (1)^2 + (2)^2}{5 - 1}$$

$$s^2 = \frac{4 + 0 + 1 + 1 + 4}{4} = \frac{10}{4} = 2.5 \text{ meses}$$

Respuesta: La varianza en los tiempos de construcción de estos proyectos es de 2.5 meses. Esto indica la medida en que los tiempos individuales de construcción

se desvían del tiempo medio de construcción.

Interpretación: Una varianza de 2.5 meses sugiere que hay una variabilidad moderada en los tiempos de construcción de estos proyectos. Comprender esta variabilidad es esencial para la planificación futura y la gestión de riesgos en proyectos de construcción.

13. Desarrollo:

Cálculo de la Media:

Media = (Suma de todas las resistencias) / (Número de proveedores)

$$\text{Media} = (30 + 35 + 32 + 34 + 33) / 5 = 164 / 5 = 32.8 \text{ MPa}.$$

Cálculo de la Varianza:

$$\text{Varianza } (s^2) = \sum \frac{(X_i - \bar{X})^2}{n - 1}$$

En este caso:

X_i : la resistencia a la compresión de cada muestra de concreto.

$\bar{X}$: la media de las resistencias a la compresión.

n : la cantidad de muestras de concreto

$$s^2 = \frac{(30 - 32.8)^2 + (35 - 32.8)^2 + (32 - 32.8)^2 + (34 - 32.8)^2 + (33 - 32.8)^2}{5 - 1}$$

$$s^2 = \frac{(-2.8)^2 + (2.2)^2 + (-0.8)^2 + (1.2)^2 + (0.2)^2}{5 - 1}$$

$$s^2 = \frac{7.84 + 4.84 + 0.64 + 1.44 + 0.04}{4} = \frac{14.8}{4} = 3.70 \text{ MPa}^2$$

Respuesta: La varianza en la resistencia a la compresión del concreto entre los proveedores es de 3.7 MPa².

Análisis: Esto indica que hay una variabilidad moderada en la resistencia del concreto suministrado por los diferentes proveedores. Aunque todos los proveedores cumplen con un estándar mínimo de calidad (suponiendo que 30 MPa es el mínimo aceptable), la variabilidad sugiere que algunos proveedores ofrecen un concreto consistentemente más resistente que otros. Este análisis ayuda a tomar decisiones informadas sobre qué proveedores seleccionar para garantizar la calidad y la seguridad en el proyecto de construcción.

14. Solución

	Xi	Xi-Xmedia	(Xi-Xmedia)²
1	10	-2,4	5,76
2	12	-0,4	0,16
3	15	2,6	6,76
4	11	-1,4	1,96
5	14	1,6	2,56
	62		17,2

El coeficiente de variación de los costos de los materiales es **14.9%**.

15. Solución

N	Xi	Xi-Xmedia	(Xi-Xmedia)²
1	120	-15	225
2	150	15	225
3	130	-5	25
4	140	5	25
5	135	0	0
	675		500

El coeficiente de variación de las alturas de los edificios es **7.4%**.

16. Solución

N	Xi	Xi-Xmedia	(Xi-Xmedia)²
1	12	-3	9
2	14	-1	1
3	16	1	1
4	18	3	9
5	15	0	0
	75		20

El coeficiente de variación de la duración de los proyectos es **13.33%**.

17. Solución

N	Xi	Xi-Xmedia	(Xi-Xmedia)²
1	20	-1,2	1,44
2	22	0,8	0,64
3	25	3,8	14,44
4	18	-3,2	10,24
5	21	-0,2	0,04
	106		26,8

El coeficiente de variación de la cantidad de concreto utilizado es **10.96%**.

18. Solución

N	Xi	Xi-Xmedia	(Xi-Xmedia) ²
1	32	-1	1
2	35	2	4
3	31	-2	4
4	33	0	0
5	34	1	1
	165		10

El coeficiente de variación de las temperaturas máximas diarias es 4.27%.

19. Solución

N	Xi	Xi-Xmedia	(Xi-Xmedia) ²
1	3,5	-0,32	0,1024
2	4	0,18	0,0324
3	3,8	-0,02	0,0004
4	4,2	0,38	0,1444
5	3,6	-0,22	0,0484
	19,1		0,328

El coeficiente de variación de los anchos de los tramos de carretera es 6.7%.

20. Solución

N	Xi	Xi-Xmedia	(Xi-Xmedia) ²
1	2	-0,3	0,09
2	2,5	0,2	0,04
3	3	0,7	0,49
4	1,8	-0,5	0,25
5	2,2	-0,1	0,01
	11,5		0,88

El coeficiente de variación de las profundidades de las fundaciones es **18.3%**.

21. Solución

N	Xi	Xi-Xmedia	(Xi-Xmedia) ²
1	30	-1,8	3,24
2	35	3,2	10,24
3	28	-3,8	14,44
4	32	0,2	0,04
5	34	2,2	4,84
	159		32,8

El coeficiente de variación de las áreas de las habitaciones es 8.05%

22. Solución

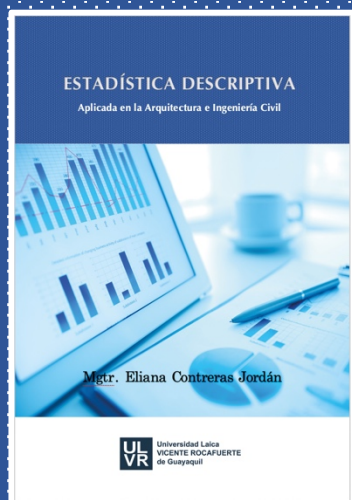
N	Xi	Xi-Xmedia	(Xi-Xmedia)²
1	80	-20	400
2	120	20	400
3	100	0	0
4	90	-10	100
5	110	10	100
	500		1000

El coeficiente de variación de los caudales de los ríos es 14.14%.

23. Solución

N	Xi	Xi-Xmedia	(Xi-Xmedia)²
1	5000	-1500	2250000
2	7000	500	250000
3	6000	-500	250000
4	8000	1500	2250000
5	6500	0	0
	32500		5000000

El coeficiente de variación de las poblaciones en las urbanizaciones es 15%.



Estadística descriptiva aplicada en la Arquitectura e Ingeniería Civil es una obra integral diseñada para formar a futuros profesionales y fortalecer la práctica de quienes ya se desempeñan en el ámbito de la construcción. Este libro ofrece un recorrido detallado por los conceptos y técnicas fundamentales de la estadística descriptiva, desde la recopilación y organización de datos hasta su análisis y representación gráfica, permitiendo transformar cifras en herramientas para la toma de decisiones precisas y fundamentadas.

A lo largo de sus cinco capítulos, el lector encontrará un enfoque práctico y riguroso que lo prepara para enfrentar los desafíos de la ingeniería civil y la arquitectura en el siglo XXI. Con ejemplos claros y aplicaciones reales, se destacan los métodos que facilitan la interpretación de la variabilidad en proyectos, optimizando recursos, mejorando la eficiencia y garantizando la calidad en cada fase del proceso constructivo.

Esta obra invita a desarrollar un pensamiento crítico y analítico que va más allá de los cálculos, impulsando la capacidad de identificar tendencias y patrones en datos complejos. Con un lenguaje accesible y una estructura coherente, *Estadística descriptiva aplicada en la Arquitectura e Ingeniería Civil* se presenta como una herramienta indispensable para quienes desean dominar el uso de la estadística en la planificación y ejecución de proyectos de infraestructura, abriendo nuevas perspectivas para la innovación y el progreso en el sector.

ISBN: 978-9942-617-16-3



**Universidad Laica
VICENTE ROCAFUERTE
de Guayaquil**