

Capítulo 2

<https://doi.org/10.62486/978-628-97230-1-4.ch02>

Los fundamentos: conceptos clave y herramientas básicas de la estadística descriptiva

Antes de comenzar a resolver ejercicios o aplicar fórmulas, es importante entender de qué estamos hablando cuando decimos "datos", "medidas" o "análisis". La estadística no es simplemente una técnica matemática; es un lenguaje para leer la realidad (Ross, 2018). Así como las palabras nos permiten describir lo que sentimos o pensamos, la estadística nos permite describir lo que observamos en el mundo. Y para hacerlo bien, primero hay que conocer sus fundamentos.

Trabajar con estadística significa, ante todo, trabajar con datos. Pero ¿qué es un dato? En términos simples, un dato es cualquier información recolectada sobre una característica específica de personas, objetos o fenómenos (Bernal Pablo, 2018). Esa información puede surgir de preguntas en una encuesta, de mediciones en un experimento o de observaciones cotidianas. Por ejemplo, si registramos la edad de cada persona en un salón de clases, cada número recogido es un dato. Y al conjunto completo de edades lo llamamos serie de datos.

Esa serie puede ser extensa, caótica o incluso confusa al principio. Aquí entra en juego la estadística descriptiva, cuya función principal es ayudarnos a organizar y comprender ese conjunto de datos sin necesidad de hacer predicciones. Es como tener una caja llena de piezas y clasificar cada una para entender cómo encajan entre sí. Como explican Moore et al. (2017), la estadística descriptiva permite resumir grandes cantidades de información en formas más manejables, para que podamos detectar patrones y diferencias de manera eficiente.

Uno de los primeros conceptos importantes es el de población. En estadística, no hablamos de población como sinónimo de país o ciudad, sino como el conjunto total de elementos que comparten una característica común y que nos interesa estudiar (Arias-Gómez et al., 2016). Puede ser un grupo de estudiantes, de celulares, de árboles en un bosque, de días lluviosos o de cualquier otra cosa que sea objeto de observación. En cambio, una muestra es un subconjunto de esa población, más pequeño, que seleccionamos para analizar con la intención de sacar conclusiones sobre el todo (Bueno, 2008). Esta relación entre población y muestra es esencial, porque en la mayoría de los casos no es posible estudiar a todos los elementos, por lo que se trabaja con una parte representativa.

A partir de esa muestra o población, podemos aplicar un censo o generar estadísticas. Un censo, según Tabutin (1997) ocurre cuando recolectamos información de todos los elementos posibles (como cuando se hace un censo nacional de habitantes). Por el contrario, una estadística es el resultado de estudiar solo una muestra, es decir, una parte. Cuando el valor calculado corresponde a toda la población, se llama parámetro; si proviene de una muestra, se llama estadístico. Esta diferencia entre parámetro y estadístico es uno de los pilares del análisis cuantitativo (Anderson et al., 2015), y es fundamental para no caer en errores al interpretar los resultados.

Los datos, una vez recolectados, suelen estar en bruto, es decir, sin ningún tipo de orden o estructura. A este conjunto se le conoce como datos en bruto, y representan el punto de partida del análisis (López et al., 2006). A medida que avanzamos, esos datos se pueden ordenar de menor a mayor, agrupar en intervalos o clasificar en categorías. Esta organización se llama distribución de frecuencias, y consiste en contar cuántas veces se repite un mismo valor o grupo de valores (Falcón y Herrera, 2005).

Ahora bien, no todos los datos son iguales. Algunos datos son cualitativos, es decir, representan características no numéricas, como el color de ojos, el género

o la nacionalidad. Otros son cuantitativos, lo que significa que expresan cantidades, como el número de hijos, la edad o el ingreso mensual. Dentro de los datos cuantitativos, además, se pueden distinguir dos tipos: los discretos, que solo toman valores enteros (como el número de estudiantes en una clase), y los continuos, que pueden asumir cualquier valor dentro de un rango (como la estatura o el peso).

Estos conceptos nos ayudan a entender qué tipo de análisis debemos hacer. Por ejemplo, si queremos analizar las edades de un grupo de personas, sabemos que estamos trabajando con una variable cuantitativa continua. Pero si estamos observando el tipo de música favorita de los estudiantes, hablamos de una variable cualitativa nominal.

Una vez que comprendemos qué tipo de datos tenemos y cómo se organizan, podemos comenzar a analizar sus propiedades. Aquí entran las medidas de tendencia central, que nos indican dónde se concentra la mayor parte de los datos, y las medidas de dispersión, que nos dicen qué tan alejados están unos valores de otros. En los siguientes capítulos aprenderemos a calcularlas, pero en este momento es importante entender para qué sirven y cómo se interpretan.

La media aritmética, comúnmente conocida como el promedio, es probablemente la medida más usada. Se calcula sumando todos los datos y dividiéndolos entre la cantidad total de valores. La media nos da una idea general del “centro” del conjunto, aunque puede verse afectada por valores extremos. Por eso, en ciertos casos conviene utilizar la mediana, que es el valor que se encuentra justo en el centro cuando los datos están ordenados. La mediana no se ve tan influida por los extremos y puede ofrecer una mejor idea de la tendencia central en distribuciones sesgadas. La moda, por su parte, es el valor que más se repite en un conjunto. Aunque muchas veces se le da menos importancia, la moda es útil para detectar comportamientos comunes o frecuentes en un grupo.

En cuanto a la dispersión, existen varias herramientas para medirla. La más conocida es la desviación estándar, que indica cuánto varían los datos respecto a la media. Si la desviación es baja, significa que la mayoría de los valores están cerca del promedio. Si es alta, los datos están más dispersos. Otra medida relevante es la varianza, que es la raíz cuadrada de la desviación estándar al cuadrado. Ambas se usan para describir la variabilidad y para identificar si existen datos muy alejados de los demás. También existen medidas como el coeficiente de variación, útil para comparar la dispersión entre conjuntos diferentes, y los cuartiles, que dividen los datos en cuatro partes iguales para analizar su distribución interna.

Este tipo de análisis se puede realizar de dos formas: usando datos no agrupados, cuando trabajamos directamente con los valores individuales, o datos agrupados, cuando organizamos los datos en intervalos o clases. Ambos enfoques tienen sus ventajas, y aprenderemos a manejarlos paso a paso.

En este capítulo hemos recorrido los principales conceptos que necesitas dominar para entender la estadística descriptiva. A partir de ahora, cada fórmula, cada ejercicio y cada análisis tendrán más sentido, porque ya conoces la base sobre la que están contruidos. No es necesario memorizar todo de inmediato, pero sí es fundamental reconocer los términos y comprender su función. Si en algún momento sientes que algún concepto se vuelve confuso, vuelve a este capítulo y repasa. La estadística, como cualquier idioma, se fortalece con la práctica y la repetición.

En el siguiente capítulo comenzaremos a aplicar lo aprendido a través de ejercicios con datos no agrupados. Verás que, una vez entendido el “idioma” de la estadística, resolver problemas se vuelve una tarea mucho más sencilla y hasta entretenida. A medida que avancemos, aprenderás no solo a calcular, sino también a interpretar, y con ello, a tomar decisiones mejor fundamentadas.

Derechos de Autor (Copyright) Guillermo Alejandro Zaragoza Alvarado 2025 ©

Este texto está protegido por una licencia Creative Commons 4.0.



Usted es libre de compartir, copiar y redistribuir el material en cualquier medio o formato, así como de adaptarlo, remezclarlo, transformarlo y crear a partir de él para cualquier propósito, incluso con fines comerciales. Sin embargo, debe cumplir con la condición de atribución, lo que significa que debe otorgar el crédito correspondiente a la obra original de manera adecuada, proporcionar un enlace a la licencia e indicar si se han realizado modificaciones. Puede hacerlo en cualquier formato razonable, pero no de manera que sugiera que cuenta con el respaldo del licenciante o que recibe algún beneficio por el uso de la obra.

[Resumen de licencia](#) – [Texto completo de la licencia](#)

ISBN: 978-628-97230-1-4

DOI: 10.62486/978-628-97230-1-4.ch02