



Chapter 09 / Capítulo 09

Evidence-based biomedicine: methodology for research, standardization, and scientific procedural aspects (Spanish Edition)

ISBN: 978-9915-704-03-6

DOI: 10.62486/978-9915-704-03-6.ch09

Pages: 52-56

©2025 The authors. This is an open access article distributed under the terms of the Creative Commons Attribution (CC BY) 4.0 License.

Statistics / Estadística

La estadística es una disciplina científica enfocada en la sistematización, recopilación, organización y presentación de datos relacionados con fenómenos caracterizados por la variabilidad o incertidumbre. Su propósito es estudiar estos fenómenos de manera metódica para inferir las leyes que los rigen y, a partir de ahí, realizar pronósticos, tomar decisiones o llegar a conclusiones. De manera general, los análisis estadísticos se categorizan en dos tipos: Estudios Descriptivos, que se centran en resumir los datos, y Estudios Inferenciales, que permiten hacer proyecciones, tomar decisiones o extraer conclusiones a partir de los datos. Los pasos fundamentales en un estudio estadístico incluyen:

1. Formulación de hipótesis.
2. Análisis descriptivo de la muestra recogida.
3. Aplicación de técnicas de inferencia para extraer conclusiones sobre la población en estudio.

9.1. Conceptualización básica

Antes de iniciar un análisis estadístico, es fundamental comprender algunos conceptos clave. Estos incluyen la población, que es el conjunto total de interés para el análisis. A menudo, la población es tan extensa que es impracticable o demasiado costoso acceder a todos sus miembros. Por ello, se recurre a una muestra, que es un subconjunto representativo de la población. La información que recopilamos de cada individuo en la muestra se describe a través de variables, que pueden ser de dos tipos:

1. Cualitativas: estas variables no son medibles numéricamente y se refieren a:
 - Atributos: son variables nominales que se dividen en diferentes categorías, como dicotómicas (hombre/mujer, soltero/casado, expuesto/no expuesto) o politómicas (grupo sanguíneo).
 - Ordinales: variables donde las categorías pueden ordenarse jerárquicamente (por ejemplo, nivel de dolor del paciente: severo, moderado, leve, ninguno; grado funcional del paciente: I, II, III, IV).
2. Cuantitativas: son variables que se pueden medir y se clasifican en:
 - Discretas: aquellas cuyo dominio es un conjunto numerable de valores y no permiten valores intermedios entre dos valores consecutivos (por ejemplo, número de hijos, número de intervenciones quirúrgicas).
 - Continuas: su dominio es continuo y pueden tomar infinitos valores entre dos puntos (altura, peso, tensión arterial).

9.2. Descripción numérica: distribución de frecuencias

La distribución de frecuencias es una herramienta que organiza y presenta las observaciones disponibles de manera ordenada.

Para distribuciones de frecuencias no agrupadas, esta se define por los siguientes elementos: los valores de la variable (x_i), la frecuencia de cada valor (n_i), la proporción en que cada valor ocurre (f_i), y la proporción acumulada hasta un valor dado (F_i).

En el caso de distribuciones de frecuencias agrupadas, que son útiles cuando la variable tiene una gran cantidad de valores distintos, los datos se organizan en intervalos. Una dificultad común al crear una tabla con datos agrupados es que el límite superior de un intervalo puede

coincidir con el límite inferior del siguiente. Para solucionar esto, se asume que cada intervalo está cerrado en su extremo inferior y abierto en el superior. Esta manera de organizar los datos facilita la identificación de patrones en las respuestas.

| L_{i-1}, L_i | n_i | f_i | F_i |
|------------------|-------|-------|-------|
| $[L_0, L_1[$ | n_1 | f_1 | F_1 |
| $[L_1, L_2[$ | n_2 | f_2 | F_2 |
| ... | ... | ... | ... |
| $[L_{n-1}, L_n[$ | n_n | f_n | F_n |

Figura 9.1. Distribución de frecuencia por medio de intervalos de confianza

9.3. Representaciones gráficas

Además de los estadísticos ya mencionados, podemos resumir datos utilizando representaciones gráficas, cuyas formas varían según el tipo de variable: para variables cualitativas, las representaciones gráficas comunes son:

- Diagrama de barras: en este diagrama, las categorías de la variable se representan en el eje vertical (ordenadas) y las frecuencias relativas en el eje horizontal (abscisas). Las alturas de las barras son proporcionales a las frecuencias, ya sean absolutas o relativas.
- Diagrama de sectores: un círculo se divide en secciones, cada una representando una categoría de la variable. El área de cada sección es proporcional a la frecuencia de la categoría que representa.

Para variables cuantitativas, en particular las continuas, se utilizan:

- Histogramas: en un histograma, los rectángulos están adyacentes, reflejando la naturaleza continua de la variable. El punto medio de cada intervalo sirve como referencia para cada rectángulo, representando la distribución de los datos en intervalos continuos.

9.4. Análisis descriptivo

Como se ha mencionado, el análisis descriptivo es útil para resumir los datos. Este resumen puede realizarse a través de tablas de frecuencias o mediante representaciones gráficas. Para resumir los datos, disponemos de ciertos estadísticos (valores numéricos derivados de la muestra), que se clasifican en dos categorías:

1. Estadísticos de centralización: estos proporcionan información sobre los valores típicos o promedios de un conjunto de datos. Entre los estadísticos de centralización que abordaremos están:

- Media: calculada como la suma de los valores dividida por el número total de observaciones. Es susceptible a la influencia de valores extremos.
- Mediana: es el valor que divide el conjunto de datos por la mitad, dejando el 50 % de las observaciones a cada lado. No se ve afectada por valores extremos.
- Moda: representa el valor o valores que aparecen con mayor frecuencia. A diferencia de la media y la mediana, puede haber más de una moda, resultando en distribuciones unimodales o multimodales.

2. Estadísticos de dispersión: estos indican cuán dispersos o concentrados están los datos en torno al valor central. Entre mayor sea la dispersión, más alejados estarán los valores del centro. Incluyen:

- **Varianza:** mide el promedio de las desviaciones al cuadrado de las observaciones respecto a la media. Sensible a valores extremos y de interpretación complicada debido a que sus unidades son cuadráticas.

- **Desviación típica:** es la raíz cuadrada de la varianza, expresada en las mismas unidades que los datos originales, y comúnmente usada para describir la variabilidad de los datos en investigaciones.

- **Coefficiente de variación de Pearson:** una medida de dispersión relativa que facilita comparar la dispersión entre diferentes distribuciones. Se calcula como la relación entre la desviación típica y la media, y se expresa generalmente en porcentaje. Por ejemplo, si la media es 10 y la desviación típica es 5, la desviación típica representa el 50 % de la media.

3. Estadísticos de Forma:

- **Asimetría:** se refiere a cómo las distribuciones pueden variar entre ser simétricas o asimétricas. En una distribución con asimetría positiva, la cola del lado derecho es más extensa, resultando en un coeficiente de asimetría positivo. Por el contrario, en una distribución con asimetría negativa, la cola del lado izquierdo es más larga, y el coeficiente será negativo. En distribuciones perfectamente simétricas, el coeficiente de asimetría es cero. En tales distribuciones, la media, la mediana y, en caso de ser única, la moda coincide.

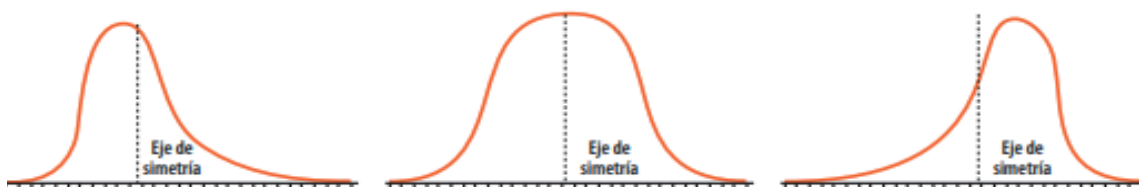


Figura 9.2. Asimetría y simetría de distribuciones

- **Curtosis:** este estadístico evalúa cuán puntiaguda o achatada es una distribución comparada con una distribución normal. Dependiendo de su valor, la distribución puede clasificarse como platicúrtica (curtosis menor que 0), mesocúrtica (curtosis igual a 0), o leptocúrtica (curtosis mayor que 0).



Figura 9.3. Apuntamiento de la distribución

4. Estadísticos de posición:

- **Cuartiles:** son tres valores que dividen un conjunto de datos, ordenado de menor a mayor o viceversa, en cuatro partes iguales. Cada parte representa el 25 % de los datos.

- **Deciles:** son nueve valores que dividen un conjunto de datos, también

ordenado de forma ascendente o descendente, en diez segmentos iguales. Cada segmento contiene el 10 % de los datos.

- Percentiles: son noventa y nueve valores que organizan un conjunto de datos, ordenado de menor a mayor o al contrario, en cien partes iguales. Cada parte abarca el 1 % de los datos.