

Módulo 4 – Inferencia Estadística

Distribución de Probabilidad

Ciencia de Datos

Objetivos de Aprendizaje



- Explicar el concepto de distribución de probabilidad.
- Reconocer las distintas distribuciones de probabilidad.
- Realizar cálculos básicos de probabilidad.

Contenido

1. Datos y variables.
2. Distribución de probabilidad.



Datos y Variables

Datos v/s Variables

¿Qué se entiende en estadística por “variable”?

Variable es cada una de las características que pueden observarse de un elemento de la muestra, es una característica de un objeto o individuo. Por ejemplo:

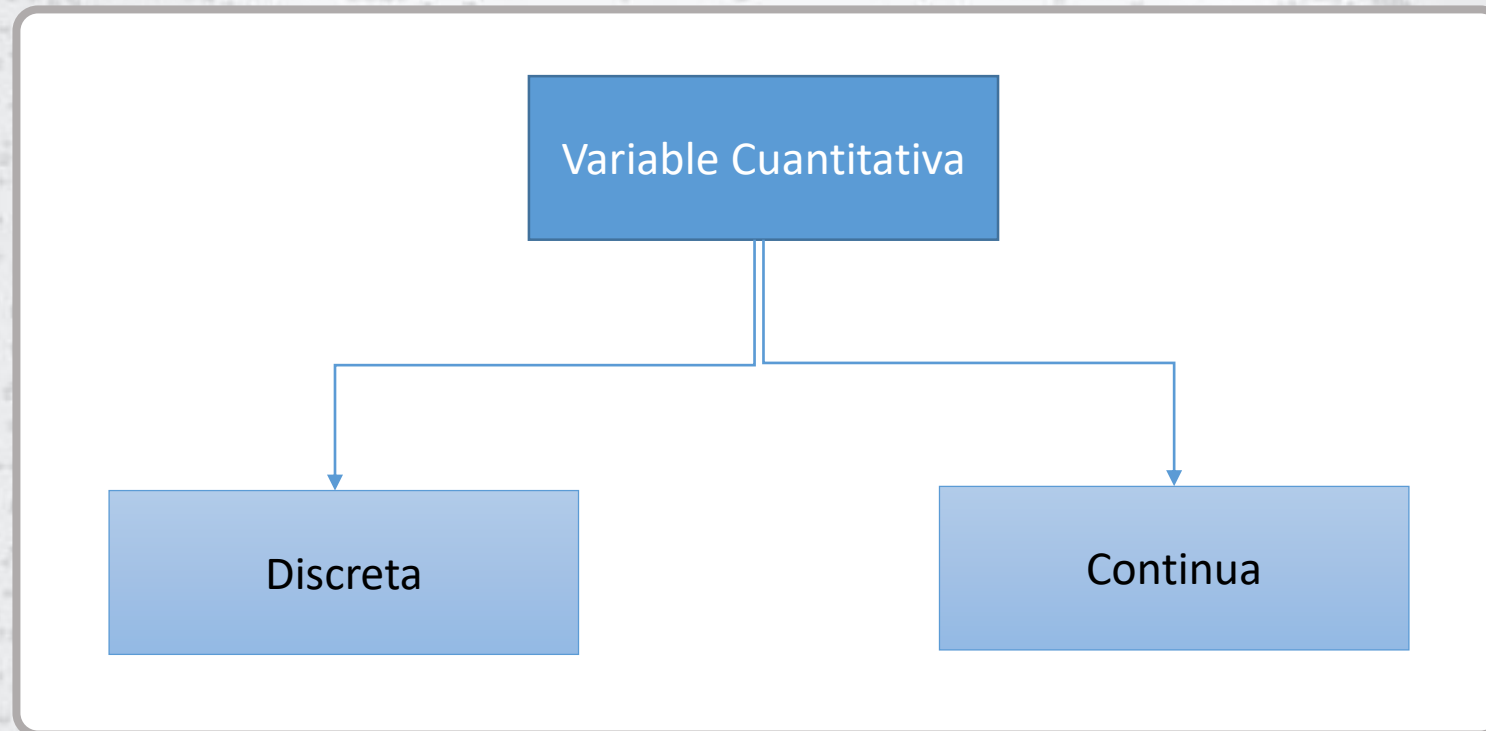
- La edad.
- El ingreso.
- Nivel de estudios.

¿Qué es un dato?

Es un valor o conjunto de valores individuales asociados con una variable.

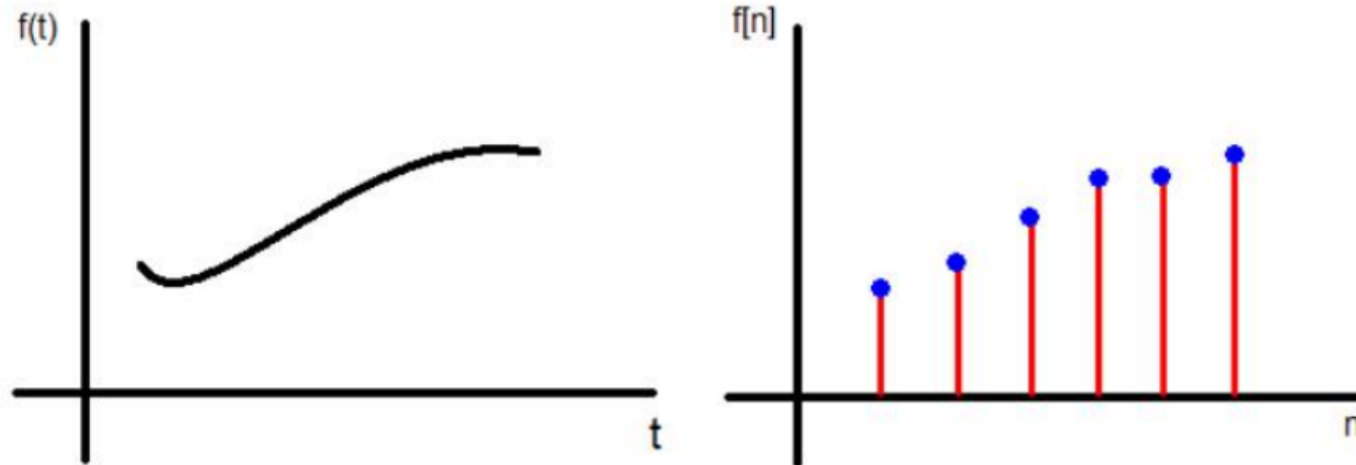
Variables Numéricas

Una variable numérica o cuantitativa tienen valores que representan cantidades. Las variables numéricas se clasifican como discretas o continuas.



Variables Numéricas

- Las **variables numéricas continuas** producen resultados que típicamente proviene de un proceso de medición, por ejemplo, la estatura o el peso de una persona, la temperatura, etc.
- Las **variables numéricas discretas** producen resultados que provienen de un proceso de conteo, por ejemplo: el número de revistas a las que una persona está suscrita o la cantidad de hijos de una persona, cantidad de piezas defectuosas, etc.



Distribución de Variables

- La información que se recoge de una o varias variables se presenta en tablas que representan la distribución de dichas variables y también, se pueden clasificar en:
- **Distribuciones unidimensionales** – Sólo recogen información sobre una característica (por ejemplo: edad de los alumnos de una clase).
 - **Distribuciones bidimensionales** – Recogen información sobre dos características de cada elemento de la población simultáneamente (por ejemplo: edad y altura de los alumnos/as de una clase) .
 - **Distribuciones multidimensionales** – Recogen información sobre tres o más características de cada elemento (por ejemplo: edad, altura y peso de los alumnos/as de una clase).

¿Qué es una Distribución de Probabilidad?

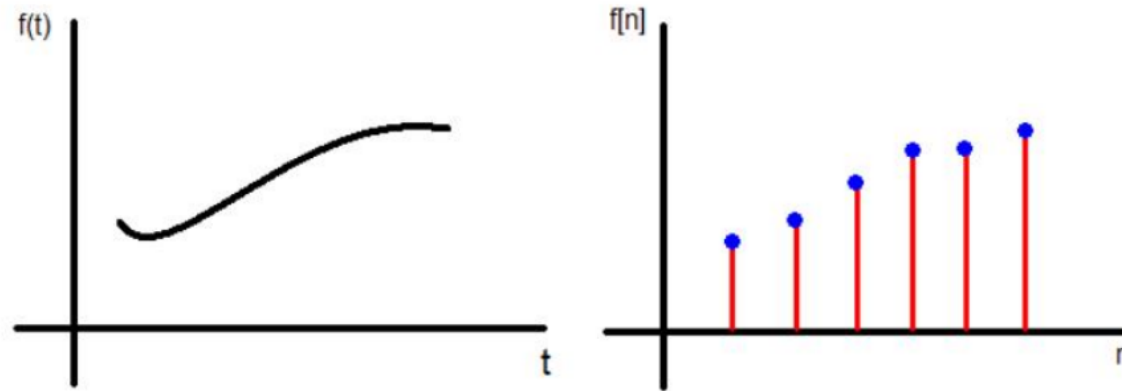
- Una distribución de probabilidad P es una función que define la probabilidad de ocurrencia de cada valor de una variable aleatoria X . Es decir, una distribución de probabilidad es una función matemática que describe las probabilidades de todos los posibles resultados de un experimento aleatorio.

$$\begin{aligned} P: \text{Img}(X) &\rightarrow [0, 1] \\ x &\mapsto P(X = x) \end{aligned}$$

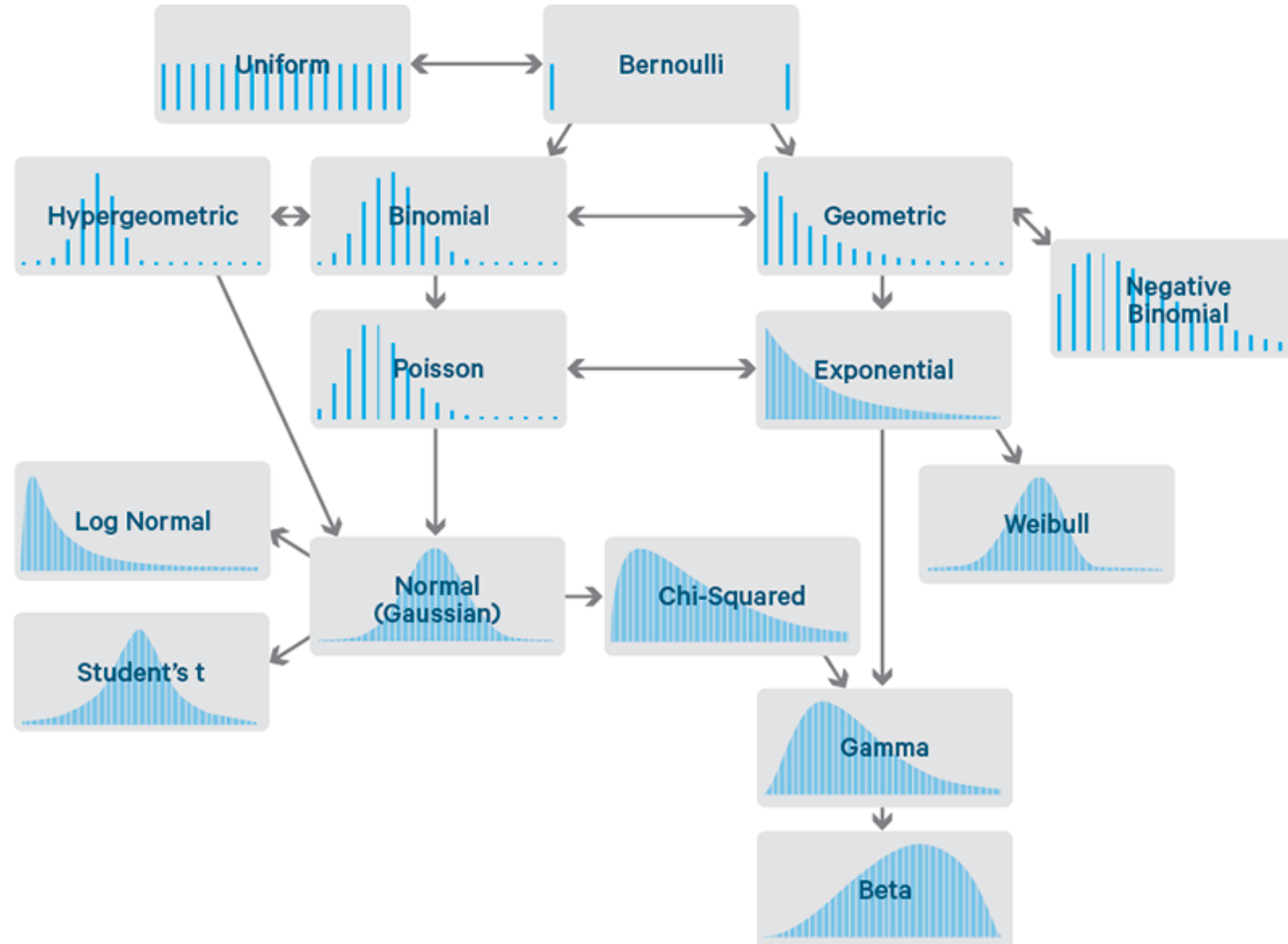
- Las distribuciones de probabilidad se usan frecuentemente en la teoría de la probabilidad y estadística, ya que sirven para calcular las probabilidades de los diferentes eventos de un espacio muestral.

Variables Numéricas

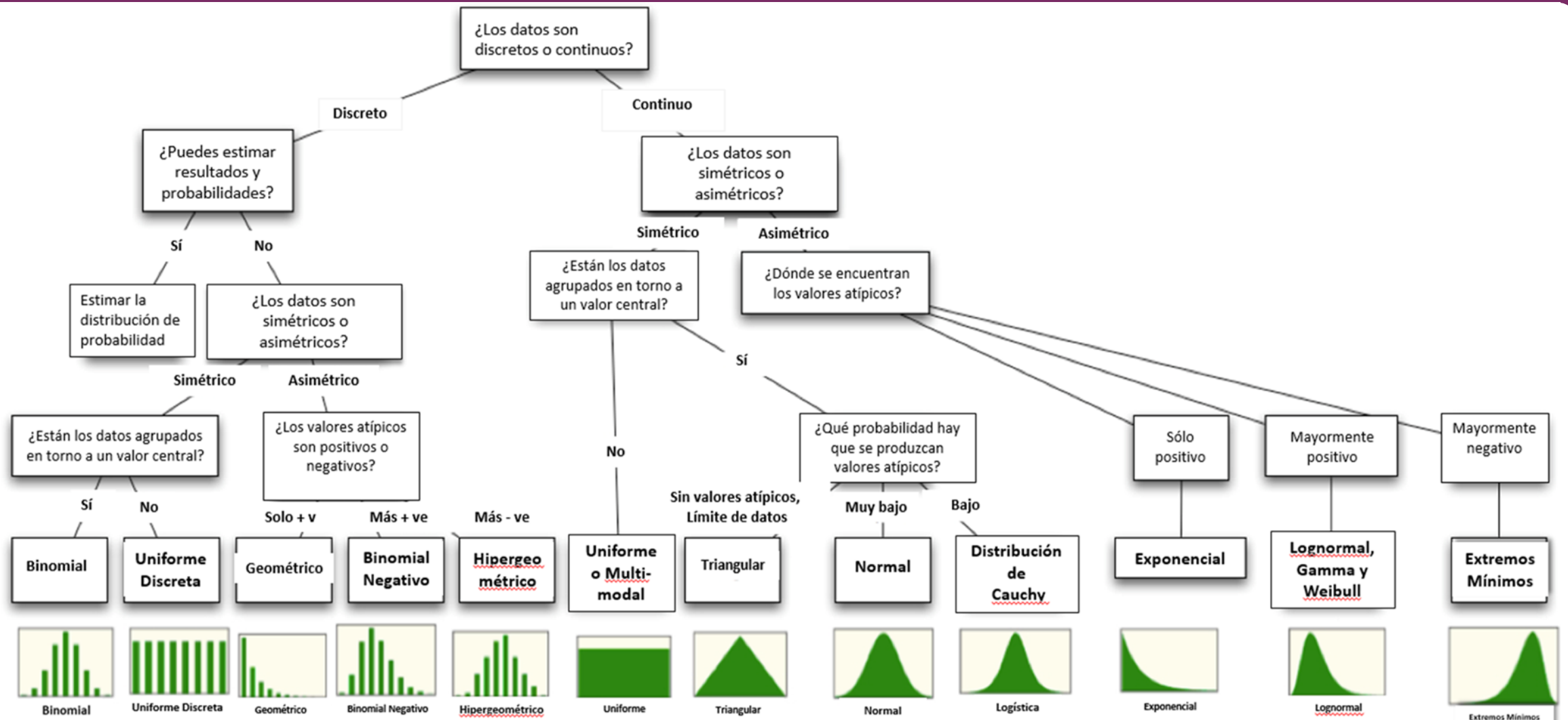
- Las distribuciones de probabilidad se pueden separar en dos grandes tipos: las distribuciones discretas y las distribuciones continuas.
- **Distribución de probabilidad discreta:** la distribución solo puede tomar un número contable de valores dentro de un intervalo. Normalmente, las distribuciones de probabilidad discretas solo pueden tomar valores enteros, es decir, que no tienen decimales.
- **Distribución de probabilidad continua:** la distribución puede tomar un número infinito de valores dentro de un intervalo. En general, las distribuciones de probabilidad continuas pueden tomar valores decimales.



Distribuciones de Probabilidad



¿Cómo elegir la mejor distribución?



Dudas y consultas

Fin presentación