

# Más allá de la regresión lineal: introducción a la regresión no paramétrica y semiparamétrica

Pamela Llop (FIQ-UNL & CONICET)

En este curso exploraremos métodos de regresión que pueden utilizarse cuando las hipótesis de linealidad impuestas por los modelos clásicos no describen adecuadamente la relación entre las variables. Comenzaremos con un breve repaso de los fundamentos de la regresión lineal simple y múltiple, analizando sus supuestos, ventajas y limitaciones.

Los modelos paramétricos poseen la ventaja de ser interpretables, computacionalmente eficientes y requerir la estimación de un número reducido de parámetros. Sin embargo, cuando la forma funcional especificada no representa correctamente la relación entre las variables, pueden producir estimaciones sesgadas y predicciones poco precisas.

Como alternativa, introduciremos los métodos de regresión no paramétrica, los cuales permiten modelar relaciones complejas sin asumir una forma funcional predeterminada para la función de regresión. Estudiaremos en detalle técnicas de suavizado y estimación no paramétrica, sus propiedades teóricas y asintóticas, así como los aspectos prácticos relacionados con la selección de parámetros de suavizado y la validación de modelos. Analizaremos las ventajas de estos métodos en términos de flexibilidad y capacidad de adaptación a la estructura de los datos, así como sus limitaciones vinculadas a la interpretabilidad y a la denominada “maldición de la dimensionalidad”.

A lo largo del curso aplicaremos estas metodologías a problemas reales provenientes de distintas áreas. Además, discutiremos extensiones de los métodos no paramétricos para datos funcionales, series temporales y datos espaciales, mostrando cómo estas herramientas pueden utilizarse en contextos modernos de análisis de datos.

Finalmente, introduciremos los modelos semiparamétricos, que combinan la interpretabilidad de los modelos paramétricos con la flexibilidad de los métodos no paramétricos. En estos modelos, una parte de la estructura se describe mediante parámetros de interés asociados a covariables explicativas, mientras que otra parte se modela de manera no paramétrica. Estudiaremos los principales modelos semiparamétricos, los algoritmos utilizados para la estimación de sus parámetros y funciones, así como sus propiedades teóricas y aplicaciones. También analizaremos extensiones a datos funcionales, series temporales y datos espaciales, destacando su utilidad en situaciones donde se busca un equilibrio entre flexibilidad, interpretación y capacidad predictiva.