

# SESSION 5 SUMMARY

Victor Miguel Terrón Macias

24/1/2021

## SESION 5. REGRESION LINEAL Y CLASIFICACIÓN

### INTRODUCCIÓN

Supongamos que nuestro trabajo consiste en aconsejar a un cliente sobre cómo mejorar las ventas de un producto particular, y el conjunto de datos con el que disponemos son datos de Publicidad que consisten en las ventas de aquel producto en 200 diferentes mercados, junto con presupuestos de publicidad para el producto en cada uno de aquellos mercados para tres medios de comunicación diferentes: TV, radio, y periódico. No es posible para nuestro cliente incrementar directamente las ventas del producto. Por otro lado, ellos pueden controlar el gasto en publicidad para cada uno de los tres medios de comunicación. Por lo tanto, si determinamos que hay una asociación entre publicidad y ventas, entonces podemos instruir a nuestro cliente para que ajuste los presupuestos de publicidad, y así indirectamente incrementar las ventas. En otras palabras, nuestro objetivo es desarrollar un modelo preciso que pueda ser usado para predecir las ventas sobre la base de los tres presupuestos de medios de comunicación. En este contexto, los presupuestos de publicidad son las variables de entrada mientras que las ventas es una variable de salida. Las variables de entrada típicamente se denotan usando el símbolo  $X$ , con un subíndice para distinguirlas. Así  $X_1$  puede ser el presupuesto para TV,  $X_2$  el presupuesto para radio, y  $X_3$  el presupuesto para periódico. Las entradas tienen diferentes nombres, tales como predictores, variables independientes, características, o a veces solo variables. La variable de salida -en este caso, las ventas- frecuentemente es llamada la variable de respuesta o dependiente, y se denota típicamente con el símbolo  $Y$ .