

Resumen del TFM_Aplicación de técnicas estadísticas en muestras desbalanceadas en *pricing*

Adán Rodal Vilas

Máster en Técnicas Estadísticas
Universidade de Santiago de Compostela

Julio de 2026

Resumen

Se hace constar que el Trabajo de Fin de Máster titulado *Aplicación de técnicas estadísticas en muestras desbalanceadas en pricing* no cuenta con autorización para su publicación en abierto ni para su difusión íntegra a través de los repositorios institucionales de la universidad.

Esta restricción responde a los compromisos de confidencialidad asumidos con la entidad colaboradora, Inditex. La memoria utiliza información interna de Zara y describe procedimientos relacionados con la fijación de los precios de la marca cuya divulgación podría revelar criterios comerciales y operativos de la compañía. Por este motivo, el presente documento expone los objetivos, el enfoque estadístico y las conclusiones generales del trabajo sin reproducir datos, reglas de negocio, ni resultados o ejemplos cuantitativos concretos.

1. Introducción y definición del problema

El trabajo se realizó en colaboración con Inditex y se centró en estudiar cómo afectan los cambios de precio a la demanda de Zara. El objetivo fue entender mejor esta relación y, al mismo tiempo, comprobar si era posible predecir la evolución de las ventas ante diferentes escenarios. De esta forma, se trató de ver hasta qué punto los datos históricos pueden ayudar en la toma de decisiones de *pricing*.

Uno de los principales problemas fue el desbalance de la muestra. No todos los tipos ni valores de cambios de precio aparecen con la misma frecuencia, ya que los datos proceden de decisiones comerciales tomadas en el pasado. Por tanto, hay situaciones para las que se dispone de mucha información y otras que aparecen pocas veces. Esta diferencia debe tenerse en cuenta tanto al ajustar el modelo como al interpretar sus resultados.

2. Datos y preprocesamiento

Para realizar el estudio se utilizaron datos históricos agregados de ventas y exposición, además de variables temporales y relacionadas con los cambios de precio. Los datos se organizaron de forma que se pudiera comparar la demanda antes y después de cada cambio. Así fue posible analizar las decisiones al nivel en el que se toman dentro de la marca, sin tener que seguir cada artículo de manera individual.

Antes de realizar el análisis fue necesario filtrar y preparar la información. Se eliminaron observaciones correspondientes a promociones y períodos de rebajas, porque en esos casos las ventas pueden responder a factores distintos de los que se querían estudiar. También se evitaron cambios de precio demasiado próximos entre sí y situaciones afectadas por circunstancias macroeconómicas extremas.

Con estos filtros se buscó obtener una muestra lo más comparable posible y separar el efecto del precio de otros factores que también afectan a las ventas. Aun así, los datos son observacionales y proceden de situaciones reales, por lo que no se puede asegurar que recojan todas las variables que intervienen en la demanda.

3. Análisis exploratorio

El primer análisis confirmó que la muestra estaba distribuida de forma desigual entre los diferentes tipos de cambios de precio. Este desbalance es importante porque el modelo cuenta con más información sobre los escenarios más frecuentes y tiene más incertidumbre en aquellos que aparecen pocas veces.

También se observó que las variables de ventas y exposición eran muy asimétricas. Por este motivo se aplicaron transformaciones logarítmicas, que ayudaron a reducir la heterocedasticidad y la influencia de las observaciones extremas. Además, se estudiaron las relaciones entre las variables para detectar posibles dependencias y decidir cuáles debían incorporarse al modelo.

Por último, se utilizaron contrastes no paramétricos, como los test de Wilcoxon y de Jonckheere-Terpstra. Sus resultados mostraron que existía una asociación entre los cambios de precio y la evolución de las ventas. Este análisis sirvió como punto de partida para el modelo y permitió detectar algunas de las principales limitaciones de los datos.

4. Metodología y resultados

Para analizar la relación entre las variables se utilizó un modelo aditivo con P-splines seleccionando el parámetro de penalización de la curvatura mediante el criterio de validación cruzada generalizada. Al utilizar las transformaciones logarítmicas, el modelo puede ser interpretado como un modelo multiplicativo en la escala original, que es el efecto esperado de los cambios de precio sobre la demanda. Además, este tipo de modelos permiten recoger relaciones no lineales entre la demanda, el cambio de precio y el resto de factores, manteniendo una gran interpretabilidad. Para comprobar su capacidad predictiva, el modelo se evaluó con datos que no se habían utilizado durante el ajuste.

Los resultados fueron adecuados y no se observaron señales claras de sobreajuste. El efecto

estimado del precio sobre la demanda fue decreciente y aproximadamente monótono. Es decir, si el resto de las variables se mantiene constante, una subida de precio se relaciona con una bajada de las unidades vendidas, mientras que una bajada de precio se relaciona con un aumento de la demanda. Además, el modelo permitió que este efecto variase según el tamaño del cambio realizado.

También se probaron diferentes escenarios de precio para estimar las unidades vendidas y los ingresos asociados a cada uno. Estas simulaciones sirven para comparar alternativas y ver las posibles consecuencias de una decisión antes de llevarla a la práctica.

No obstante, el modelo no permite diferenciar los distintos comportamientos que pueden existir entre mercados o tipos de producto. Por este motivo, sus resultados deben complementarse con el conocimiento de negocio y con la valoración de los departamentos responsables de las decisiones comerciales.

5. Conclusiones

En este trabajo se ha visto que las técnicas estadísticas utilizadas permiten estudiar la relación entre precio y demanda. El modelo ayuda a interpretar el efecto del precio y ofrece predicciones útiles para comparar escenarios, cumpliendo así los dos objetivos principales del proyecto.