

Examen - Semestre Primavera 2020

6 de Enero, 2021

Por favor realice el siguiente compromiso. Yo _____ responderé las preguntas a continuación en forma individual. Por lo tanto, no me comunicaré con otras personas en nada relacionado al contenido de esta evaluación.

P1. COVID-19 (Ventas Web)

Producto de la pandemia producida por el coronavirus el modo de comprar se ha concentrado principalmente en ventas online. El cambio del comportamiento de los clientes ha dejado obsoletos los modelos de predicción de ventas que ayudaban a definir los niveles de stock. En este contexto, se te ha pedido desarrollar un nuevo modelo para estimar las ventas del canal web. Para ello la empresa cuenta con una base de datos con los siguientes variables:

Variable	Descripción
SKU	Stock Keeping Unit
Categoría	Indica el tipo de producto (ej. Vestuario hombre)
Subcategoría	Indica el tipo de producto (ej. Polera piqué)
Semana	Indica en que semana del año se realizó la compra
Región	Indica la región de la dirección de despacho (o retiro)
Comuna	Indica la comuna de la dirección de despacho (o retiro)
Población	Indica la población de la comuna
Semanas en cuarentena	Indica el número de semanas seguidas que lleva la comuna en cuarentena
Venta total	Monto en pesos para el SKU indicado
Cantidad	Número de productos vendidos del SKU indicado

Tabla 1: Datos de venta online.

- (1.5 puntos) Formular un modelo de regresión lineal que permita a la empresa estudiar el nivel de ventas web de cada producto. El modelo debe considerar los siguientes factores:
 - Un efecto fijo por comuna y sub-categoría.
 - Que el nivel de ventas está directamente relacionado con la población de la comuna.
 - Que el cuarentenas prolongadas tienen un efecto en las ventas online.
- En la siguiente tabla se muestran algunos de los resultados obtenidos al estimar el modelo anterior, en donde f corresponde a una transformación que permite ajustar mejor los datos disponibles.
 - (1.0 puntos) Argumente si le parece razonable incluir efectos fijos a nivel de comuna y subcategoría.
 - (1.0 puntos) Discuta qué función f le parece razonable usar para capturar el efecto de las cuarentenas e interprete el valor de la tabla de acuerdo a su elección de f .
- (1.5 puntos) Plantee un modelo probabilístico que combine heterogeneidad observable y no observable para contar cuantas unidades de cada producto se venderán semanalmente a través del sitio web. Escriba la log-verosimilitud correspondiente.
- (1.0 puntos) Discuta los desafíos de implementar los modelos anteriores en una organización e indique tres acciones que recomendaría implementar tempranamente para *iniciar* el proceso de desarrollo.

Variable	Coefficiente	p-value
β_0	-4.522	0.0789
Aysen	-12.289	0.0821
Algarrobo	-4.284	0.3441
Alhue	-10.232	0.4151
Alto Biobio	-7.231	0.2891
...	...	...
Camisa	0.4692	0.1423
Jeans	3.6679	0.0346
Short	2.1972	0.0720
$f(\text{Semanas en cuarentena})$	8.011	0.0023

P2. COVID-19 (visita a tienda)

La pandemia del COVID-19 ha traído importantes cambios en los patrones de elección de las tiendas en que se realizan las compras de los clientes. Para entender esos cambios, se ha configurado un panel de N clientes que registra sus visitas a J posibles tiendas en las últimas T semanas. El panel registra la variable y_{njt} que toma el valor 1 si el cliente n visita la tienda j en la semana t (0 en caso contrario) y un conjunto de variables explicativas como se describe en la Tabla 2.

Variable	Descripción
RUBRO $_j$	Rubro al que pertenece la tienda j (e.g. Supermercado, Tienda de especialidad, etc.)
INGRESO $_n$	Ingreso promedio comuna de residencia cliente n
DISTANCIA $_{nj}$	Distancia desde el hogar el cliente n a la tienda j
FASE $_{nt}$	Fase deconfinamiento del cliente n en la semana t
PRECIOS $_{jt}$	Índice de precios de la tienda i en la semana j
PROMO $_{jt}$	1 si la tienda j ofrece descuentos de precio en la semana t

Tabla 2: Datos panel de visita a tiendas

- (2.0 puntos) Escriba la utilidad para un modelo de elección discreta para describir si un cliente n visita o no una cada tienda j en cada semana t , que considere que:
 - Hay un efecto fijo por semana y otro por rubro.
 - Distintas fases de confinamiento implican distintas más o menos visitas a tiendas.
 - La distancia afecta de manera no lineal (discuta qué forma funcional podría ser adecuada).
 - El efecto de los precios y las promociones depende del nivel de ingresos de la comuna.
- (2.0 puntos) Suponga un modelo logit binario en que la utilidad de visitar una tienda viene dado por:

$$v_{njt} = \beta_0 + \beta_1 \text{Distancia}_{nj} + \beta_2 \text{Promo}_{jt} \cdot 1_{\text{Fase}_{nt} \in \{III, IV\}}$$

Los datos que muestra la Tabla 3 describen a un cliente que podría haber asistido a dos tiendas en la semana 1, pero solo asiste a una. Escriba explícitamente la log-likelihood para este caso.

Tienda	Semana	Visita	Fase	Rubro	Precios	Promo	Distancia
1	1	0	II	Supermercado	\$100	SI	5,5 km
2	1	1	II	Ferretería	\$200	NO	47,4 km

Tabla 3: Datos de un cliente del panel para la primera semana.

- (2.0 puntos) Suponga que el gobierno instaure una política en que cada ciudadano (y por tanto cada cliente del panel), puede salir solo una vez en la semana y que queremos estudiar la elección de tienda en Combarbalá, donde solo el panel de datos solo considera tres tiendas principales. Suponga que se estima

un modelo probit sobre los datos y se encuentra una matriz varianza-covarianza (desnormalizada) como sigue. Interprete los resultados.

$$\Sigma = \begin{bmatrix} 0.198 & 0.533 & 0 \\ & 0.360 & 0.258 \\ & & 0.852 \end{bmatrix}$$

Hint: En este modelo hemos asumido que todos los panelistas salen una vez a la semana y que por tanto la elección corresponde a determinar qué tienda van a visitar.

!BUENA SUERTE!