

Auxiliar Repaso C2: Modelos Estructurales y Logit

Pregunta 1.

Suponga que observa las unidades compradas por N clientes en un periodo de largo T , de modo que n_k clientes compra k unidades en todo el periodo $k \in \{0,1,2,3, \dots\}$. Análisis preliminares sugieren que la data queda bien descrita por un modelo de Poisson simple sin heterogeneidad, excepto para los clientes que no compran los que son subestimados por el modelo. Es por esto por lo que se propone usar un modelo de Poisson (λ) con ceros inflados (ZIP). **Este modelo es idéntico al de Poisson, pero con una probabilidad mayor para el caso de cero ocurrencias (γ)**. Escriba la log-verosimilitud para este modelo. Explique, ¿Cuáles serían los pasos siguientes para encontrar el estimado de Máxima Verosimilitud de forma analítica?

Pregunta 1 PAUTA.

Sea X el número de unidades compradas. Entonces, la log-verosimilitud depende de los parámetros λ , el parámetro del modelo base de Poisson y de γ , la probabilidad de que un cliente compre 0 unidades del producto. Entonces, la log-verosimilitud viene dada por:

$$\begin{aligned} LL(\lambda, \gamma) &= n_0 \ln(\gamma) + \sum_{k=1}^{\infty} n_k \ln(\mathbb{P}(X = k|\lambda)) \\ &= n_0 \ln(\gamma) + \sum_{k=1}^{\infty} n_k \ln\left(\frac{\lambda^k e^{-\lambda}}{k!}\right) \end{aligned}$$

Dado que ya tenemos planteada la función de verosimilitud, los pasos a seguir son:

- Calcular la función de log-verosimilitud, dada las propiedades matemáticas del Logaritmo.
- Aplicar condición de primero orden sobre la log-verosimilitud, con respecto al parámetro que se quiere estimar, es decir, con respecto a λ, γ .
- Despejar el estimador máximo verosímil de λ, γ .

Pregunta 2.

Se quiere estimar la probabilidad de comprar una determinada marca de yogurt en un supermercado. Se cuenta con la información de si un cliente compra o no compra el yogurt, el precio del yogurt al momento de la decisión de compra, el precio de la competencia al momento de la compra, un registro en variable binaria si el yogurt estaba en promoción durante la compra e información en variable binaria si es que la competencia estaba en promoción durante la compra.

- a) ¿Qué modelo estudiado nos permite estimar la probabilidad de elección de marca?
- b) Plantee la probabilidad de elección de marca e indique los pasos a seguir para obtener resultados de manera analítica.

Pregunta 2 PAUTA.

- a) Dada que queremos estudiar la probabilidad de elección de una marca en una categoría específica, podemos notar que las marcas discretas, por ende, necesitamos un modelo de elección discreta. Para esto, plantearemos un modelo *Logit* para estimar la probabilidad de elección.
- b) Para poder definir la probabilidad del modelo *Logit* para el problema planteado, debemos darle nombre a las variables:
- ***Y_{ni}***: Toma el valor 1 si el cliente *n* elige la marca *i* y toma el valor 0 en cualquier otra ocasión.
 - ***featyog_i***: Toma el valor 1 si el yogurt *i* estaba en promoción. *priceyog_i*=precio del yogurt *i*.
 - ***featcomp_i***: Toma el valor 1 si el yogurt de la competencia estaba en promoción.
 - ***pricecomp_i***: Precio del yogurt de la competencia estaba en promoción.

Ahora sí, podemos definir la utilidad que le reporta al cliente *n* escoger la marca *i*, como u_{ni} :

$$u_{ni} = \beta_0 + \beta_1 \text{featyog}_i + \beta_2 \text{priceyog}_i + \beta_3 \text{featcomp}_i + \beta_4 \text{pricecomp}_i$$

Luego, podemos plantear la probabilidad de elección de marca de yogurt por cada cliente.

$$\mathbb{P}(Y_{ni} = 1 | \beta) = \frac{e^{u_{ni}}}{1 + e^{u_{ni}}}$$

Planteada la probabilidad, para resolverla utilizamos el método de Máxima Verosimilitud, en particular, debemos maximizar la Log-Verosimilitud variando los parámetros β_i de la utilidad. Paralelo a encontrar los parámetros desconocidos, debemos definir una probabilidad de corte para poder clasificar las elecciones de los distintos clientes. Posterior a eso, debemos analizar los resultados.