Auxiliar 4

Modelos de Elección Discreta

Profesor: Daniel Olcay
Auxiliares: Joaquín Cisternas, Gabriela Mora, Andrés Olivares

Resumen

Usamos Modelos de Elección Discreta para cuantificar y predecir las preferencias de los individuos al momento de tener que elegir una alternativa de varias. Ya sea eligiendo una marca, un medio de transporte, o candidatos políticos, siempre nos enfrentamos a elecciones. El objetivo de los modelos de Elección Discreta es entender cómo las personas eligen en distintas ocasiones.

Sea el individuo $n \in N$, elije la alternativa $i \in I$ si este maximiza su utilidad $U_{ni} = V_{ni} + \epsilon_{ni}$, donde V_{ni} es el valor observable que le entrega al individuo i la alternativa n, y ϵ_{ni} el error de medición para ese valor. Para cuantificar el comportamiento de las personas/organizaciones queremos calcular la probabilidad de que n elija i : P_{ni}.

$$\begin{aligned} P_{ni} &= P(U_{ni} \geq U_{nj} \forall j \neq i) = P(\epsilon_{nj} \leq \epsilon_{ni} + V_{ni} - V_{nj} \forall j \neq i) \\ &= \int_{\epsilon_{ni}} P(\epsilon_{nj} \leq \epsilon_{ni} + V_{ni} - V_{nj} \forall j \neq i | \epsilon_{ni}) f(\epsilon_{ni}) d\epsilon_{ni} \end{aligned}$$

Los modelos que veremos trabajan en cómo distribuye ϵ_{ni}, por lo que nos interesa saber el efecto marginal de las variables observadas en cada elección. De hecho, como las decisiones surgen de comparar la utilidad entre las distintas alternativas, en verdad nos interesa saber el efecto marginal de los **cambios** en las variables observadas. Esto significa que queremos conocer qué provoca el cambio $V_{ni} - V_{nj}$

A partir de una base de datos con registros de elecciones $Y_n = i$ a partir de variables explicativas x_{ni}, queremos utilizar esta información para comprender el valor observable del producto en las personas. De este modo, el comportamiento de los datos se acerque a:

$$U_{ni} = \alpha_i + \beta_{ni} x_{ni} + \epsilon_{ni}$$

Luego lo que queremos encontrar es α_i, β_n tal que modelen nuestros datos. Esto se obtiene a través de máxima verosimilitud. Sea $Y_{ni} = 1$ si el individuo n escogió i (0 de lo contrario) y sea $\beta = \alpha_i + \beta_n$, buscamos el parámetro:

$$\hat{\beta} = argmax LL(\beta) = \sum_n \sum_i Y_{ni} ln(P_{ni})$$

Modelo Logit

Asumiendo que cada ϵ_{ni} distribuye independientemente a valores extremos tipo I (o poseen una distribución de Gumbel(0,1)):

$$f(\epsilon_{ni}) = e^{-\epsilon_{ni}} e^{-e^{-\epsilon_{ni}}}$$
$$F(\epsilon_{ni}) = e^{-e^{-\epsilon_{ni}}}$$

Se obtiene de lo anterior que:

$$P_{ni} = \int_{\epsilon_{ni}} \Pi_{j \neq i} F(\epsilon_{ni} \leq \epsilon_{ni} + V_{ni} - V_{nj}) f(\epsilon_{ni}) d\epsilon_{ni} = \frac{e^{v_{ni}}}{\sum_j e^{v_{nj}}}$$

Luego, queremos encontrar el parámetro β que maximice la verosimilitud.

$$LL(\beta) = \sum_n \sum_i Y_{ni} ln(P_{ni})$$

Luego se quiere:

$$\frac{\partial LL(\beta)}{\partial \beta} = \sum_n \sum_i (Y_{ni} - P_{ni}) x_{ni}$$

Modelo Probit

Asumiendo que ϵ_n distribuye como una normal multivariada centrada en 0:

$$f(\epsilon_n) = \frac{1}{(2\pi)^{\frac{I}{2}} |\sum|^{\frac{1}{2}}} e^{-\frac{1}{2}\epsilon_n^T \sum^{-1} \epsilon_n}$$

Donde $\sum$ es la matriz de varianza-covarianza entre los errores de cada alternativa:

$$\sum = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \cdot & \sigma_{22} & \sigma_{23} \\ \cdot & \cdot & \sigma_{33} \end{bmatrix} \tag{1}$$

Como estamos comparando utilidades entre ellos, sólo nos importa la diferencia de sus errores y su covarianza. De los $\frac{I(I+1)}{2}$ basta conocer $\frac{I(I-1)}{2} - 1$ de ellos:

$$\hat{\sum} = \begin{bmatrix} 1 & \hat{\sigma_{23}} \\ \cdot & \hat{\sigma_{33}} \end{bmatrix} \tag{2}$$

Lamentablemente, el valor exacto de P_{ni} no se puede calcular. Afortunadamente se puede aproximar, considerando R muestras $\epsilon_{ni}^r \sim N(0, \sum)$:

$$\hat{P}_{ni} \frac{1}{R} \sum_{r=1}^{R} 1_{[\epsilon nj^r - \epsilon_{ni}^r \leq v_{ni} - v_{nj} \forall j \neq i]}$$

Y encontrar:

$$\hat{\theta} = argmaxLL(\theta) = \sum_n \sum_i Y_{ni} ln(\hat{P_{ni}})$$

P1.- Pizza a Pieza es un servicio de pizzas a domicilio que lleva varios años operando en el sector sur de la capital. Los administradores de la tienda quieren estudiar el comportamiento de compra de sus clientes. Para eso, se proponen calibrar modelos de elección discreta usando la una base de clientes registrados en la compañía y que hayan comprado al menos una vez en los últimos 6 meses. La base de datos está compuesta por 763 clientes de los que se conoce su número telefónico, la edad y género del jefe de hogar y si han comprado en cada semana t durante los últimos 6 meses de operaci´on ($y_{nt} = 1$ si el cliente n compró en la semana t). Además, para cada semana se conoce el precio de lista de las pizzas (que por políticas de la empresa es el mismo independiente de la variedad de la pizza) y si se repartieron volantes con publicidad.

a Suponga que la utilidad que un cliente n deriva por comprar en un semana t puede describirse como $u_{nt} = \beta_0 x_{nt} + \epsilon_{nt}$. Si los términos ϵ_{nt} están independiente e idénticamente distribuidos valor extremo tipo I, escriba la log-verosimilitud de un modelo que describa las compras de la base de clientes.

b Al revisar la data, los analistas se dan cuenta que en los últimos 6 meses de operación, los precios de las pizzas ha sido exactamente los mismos. Es posible estimar la sensibilidad al precio de la demanda con datos de esta naturaleza? Si no fuera posible, describa que otra información podría recolectarse para hacer la estimación.

c Después de explorar lo datos se ha acordado estimar dos modelos. El primero (M1) considera solo un intercepto y una variable de lealtad que permita diferenciar de aquellos clientes que compran mucho de aquellos que compran menos. El segundo, junto con las variables anteriores consideran además un coeficiente de precio, otro de promoción y finalmente el efecto de la publicidad (M2). Los estimadores máximo verosímiles de estos dos modelos se encuentran disponibles en la siguiente tabla:

Parámetros	M1	M2
Intercepto	5.23	5.54
Lealtad	3.17	X
Precio	-	-3.25
Promoción	-	4.73
Publicidad	-	0.04
ρ^2	0.17	Y
BIC	3145	Z

Figura 1: Estimaciones modelos alternativos Pizza a Pizza

Discuta si valores de X, Y y Z debieran ser, mayores o menores que los estimadores m´aximo verosimiles del M1.

P2.- (**Propuesto**) Suponga que un analista propone usar un modelo logit para estudiar el comportamiento de compra de 5 marcas de margarina las que pueden ofrecerse en formato pan o pote.

a Escriba la función de utilidad si los clientes tienen preferencias dependientes exclusivamente de precio y formato.

b Escriba la función de utilidad si los clientes tienen preferencias dependientes de precio, marca y formato.

c Escriba la funcióon de utilidad si la influencia que el precio tiene en la utilidad del cliente n depende del nivel de ingreso ING_n.