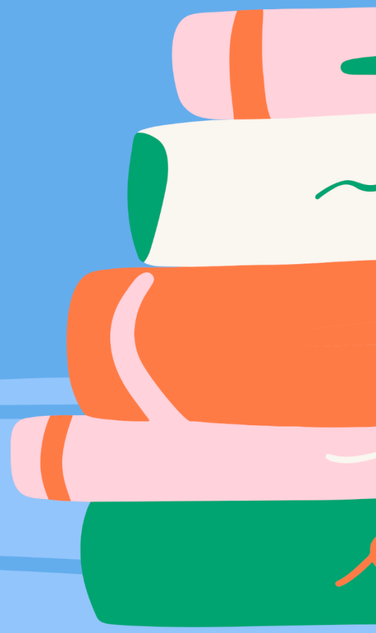
INGENIERÍA INDUSTRIAL
UNIVERSIDAD DE CHILE

Modelos de elección discreta

Otoño 2021 IN5602



¿Cuándo decidimos usarlos?

Cuando queremos describir las decisiones de los agentes entre sus alternativas

Características

01 Mutuamente excluyentes

02 Finitos

03 Exhaustivos

¿Cuál es el objetivo?

Buscamos describrir la probabilidad del individuo de elegir una de las opciones (buscan una alternativa donde la utilidad es mayor q otra)

$$P_{i,j} = Prob(U_{i,j} > U_{i,k} \ \forall \ j \neq k)$$

$$U_{i,j} = V_{i,j} + \varepsilon_{i,j}$$

$$V_{i,j} = \beta_1 X_j + \delta_i S_i$$

$$V_{i,j} = \beta_1 Caracteristica1_j + \delta_j Edad_i$$

$$V_{i,k} = \beta_1 Caracteristica2_k + \delta_k Edad_i$$

Lo único que importa es la diferencia de utilidades

Ojo con las variables sociodemográficas

Lo único que importa es la diferencia de utilidades

$$U_{i,j} = \beta_1 Caracteristica1_j + \delta_2 Edad_i + \varepsilon_{i,j}$$

$$U_{i,k} = \beta_1 Caracteristica2_k + \delta_2 Edad_i + \varepsilon_{i,k}$$

$$U_{i,j} - U_{i,k} = \beta_1 \left(Caracteristica1_j - Caracteristica2_k\right) + \delta_2 (Edad_i - Edad_i) + \varepsilon_{i,j} - \varepsilon_{i,k}$$

Lo único que importa es la diferencia de utilidades

$$U_{i,j} = \beta_1 Caracteristica1_j + \delta_2 Edad_i + \varepsilon_{i,j}$$

$$U_{i,k} = \beta_1 Caracteristica2_k + \delta_2 Edad_i + \varepsilon_{i,k}$$

$$U_{i,j} - U_{i,k} = \beta_1 \left(Caracteristica1_j - Caracteristica2_k\right) + \delta_2 (Edad_i - \cancel{Edad_i}^{\,0}) + \varepsilon_{i,j} - \varepsilon_{i,k}$$

Ojo con las variables sociodemográficas

$$U_{i,j} = \beta_1 Caracteristica1_{i,j} + \delta_j Edad_i + \varepsilon_{i,j}$$

$$U_{i,k} = \beta_1 Caracteristica2_{i,k} + \delta_k Edad_i + \varepsilon_{i,k}$$

$$U_{i,j} - U_{i,k} = \beta_1 \left(Caracteristica1_{i,j} - Caracteristica2_{i,k}\right) + \left(\delta_j - \delta_k\right) Edad_i + \varepsilon_{i,j} - \varepsilon_{i,k}$$

Lo único que importa es la diferencia de utilidades

$$U_{i,j} - U_{i,k} = \beta_1\left(Caracteristica1_j - Caracteristica2_{i,k}\right) + \left(\delta_j - \delta_k\right)Edad_i + \varepsilon_{i,j} - \varepsilon_{i,k}$$

¿Cuándo se justifica incluir interceptos en el modelo?

Modelo de probabilidad

$$P_{n,j} = Prob(U_{n,j} - U_{n,i} > 0 \ \forall \ j \neq i)$$

$$= Prob(V_{n,j} + \varepsilon_{n,j} > V_{n,i} + \varepsilon_{n,i} \forall \ j \neq i)$$

$$= Prob(\varepsilon_{n,j} - \varepsilon_{n,i} > V_{n,i} - V_{n,j} \forall \ j \neq i)$$

Logit

Probit

Logit

$$P_{i,j} = Prob(U_{n,j} > U_{n,i} \ \forall\, j \neq i)$$

$$= Prob(V_{n,j} + \varepsilon_{n,j} > V_{n,i} + \varepsilon_{n,i} \forall\, j \neq i)$$

$$= Prob(\varepsilon_{n,j} - \varepsilon_{n,i} > V_{n,i} - V_{n,j} \forall\, j \neq i)$$

$$= \frac{e^{v_{i,j}}}{\sum_k e^{v_{i,k}}}$$

Gumbel o tipo I valor extremo

Probit

$$P_{i,j} = Prob(U_{n,j} > U_{n,i} \ \forall\, j \neq i)$$

$$= Prob(V_{n,j} + \varepsilon_{n,j} > V_{n,i} + \varepsilon_{n,i} \forall\, j \neq i)$$

$$= Prob(\varepsilon_{n,j} - \varepsilon_{n,i} > V_{n,i} - V_{n,j} \forall\, j \neq i)$$

$$\approx \frac{1}{R} \sum_{r=1}^{R} 1_{[\varepsilon_{n,j} - \varepsilon_{n,i} > V_{n,i} - V_{n,j} \forall\, j \neq i]} \quad \{\varepsilon_n^r\}_{r=1}^{R} \; iid, N(0, \Sigma)$$

My Progress
My Bookmarks
Organizations
Career Tracks
Skill Tracks
Courses
Practice
Projects
Assessments
Live Events

INTERACTIVE COURSE

Choice Modeling for Marketing in R

Continue Course | Bookmark

4 hours | 17 Videos | 54 Exercises | 4,239 Participants | 4,100 XP

Course Description

People make choices everyday. They choose products like orange juice or a car, decide who to vote for, and choose how to get to work. Marketers, retailers, product designers, political scientists, transportation planners, sociologists, and many others want to understand what drives these choices. Choice models predict what people will choose as a function of the features of the options available and can be used to make important product design decisions. This course will teach you how to organize choice data, estimate choice models in R and present findings. This course covers both analyses of observed real-world choices and the survey-based approach called conjoint analysis.

This course is part of these tracks:

Marketing Analytics with R

Elea McDonnell Feit

1 Quickstart Guide FREE 100%

Anexo

WIDE

Dia de pedido	Manzanas	Bananas	Peras
Jueves	1	3	5
Viernes	2	2	1
Sabado	10	1	0
Domingo	0	2	5

LONG

Dia de pedido	Fruta	Cantidad
Jueves	Manzana	1
Jueves	Banana	3
Jueves	Peras	5
Viernes	Manzana	2
Viernes	Banana	2
Viernes	Peras	1
Sabado	Manzana	10
Sabado	Banana	1
Sabado	Peras	0
Domingo	Manzana	0
Domingo	Banana	2
Domingo	Peras	5

Anexo

Recall (también se le llama sensitivity) $= \frac{\text{True Positive}}{\text{Predicted Results}}$ or $\frac{\text{True Positive}}{\text{True Positive + False Negative}}$

Accuracy $= \frac{\text{True Positive + True Negative}}{\text{Total}}$

Especificity $= \frac{\text{True Negatives}}{\text{True Negatives + False Positives}}$