

Introduction aux choix discrets

Fabian Bastin
Département d'Informatique et de Recherche
Opérationnelle
Université de Montréal
`bastin@iro.umontreal.ca`

Avril 2011

Qui suis-je ?

Nom : Fabian Bastin

Formation : licence en sciences mathématiques aux FUNDP (1999), coctorat en mathématiques (FUNDP, 2004), séjours post-doctoraux à Londres et à Toulouse, professeur adjoint à l'Université de Montréal depuis janvier 2007.

Page web : <http://www.iro.umontreal.ca/~bastin>

Courriel : bastin@iro.umontreal.ca

Références principales

Livres :

- Kenneth Train, "Discrete Choice Methods with Simulation". Cambridge University Press, New York, USA, 2003.
Disponible à l'adresse `http://elsa.berkeley.edu/books/choice2.html`.
- Moshe Ben-Akiva, "Discrete Choice Analysis : Theory and Application to Travel Demand", The MIT Press, Boston, USA, 1985.
- Yosef Sheffi, "Urban Transportation Networks", Prentice-Hall, Englewood Cliffs, New Jersey, USA, 1985.
Disponible à l'adresse `http://web.mit.edu/sheffi/www/selectedMedia/sheffi_urban_trans_networks.pdf`.

Journaux principaux : Transportation Science, Transportation Research B, Journal of Choice Modeling.

Remerciement spécial pour Cinzia Cirillo, assistant professor à l'Université du Maryland, pour les contributions à ce cours.

Les choix discrets pour les nuls

La vie est pleine de questions, comme l'illustre par exemple le Guide du Voyageur Intergalactique. Selon ce guide, "l'Histoire de chaque Civilisation Galactique majeure tend à passer au travers de trois phases distinctes et reconnaissables, qui sont

- la Survie,
- l'Investigation,
- la Sophistication,

plus communément connues comme les phases du "Comment", du "Pourquoi" et du "Où".

Exemple :

- **Première phase** : *Comment allons-nous manger ?*
- **Deuxième phase** : *Pourquoi mangeons-nous ?*
- **Troisième phase** : *Où allons-nous dîner ?*

Supposons que nous sommes dans la troisième phase. Par conséquent, la question est :

Où allons-nous dîner ?

Différents facteurs vont influencer les réponses individuelles :

- argent disponible ;
- temps disponible ;
- distance ;
- goûts personnels (hamburger ou nouilles ?) ;
- en régime, ou lendemain du bunker ?
- ...

Un agent (personne, firme, preneur de décision) fait face à un choix, ou une série de choix au cours du temps, parmi un ensemble fini d'options.

Nous voulons observer le processus comportemental menant au choix de l'agent.

Le résultat de la situation dans n'importe quelle situation donnée est noté par y . Quelle forme aura ce y ?

Hypothèse économétrique : **les individus cherchent à maximiser leur utilité.**

En d'autres termes, on suppose que nous pouvons capturer les préférences individuelles au moyen d'une mesure d'attraction ou d'utilité associée à chaque alternative.

L'utilité est fonction

- 1 des attributs de l'alternatives ;
- 2 des caractéristiques du preneur de décision.

L'individu sélectionne alors l'alternative qui possède l'utilité maximale. Si on connaissait tout, le choix de l'agent serait complètement déterminé. . .

Modèles d'utilités aléatoires

L'utilité est connue par le preneur de décision, mais pas par le chercheur.

Puisque tout ne peut pas être observé, on décomposera une utilité entre une partie observée, déterministe, et une partie aléatoire.

Puisque les utilités sont aléatoires, nous pourrions seulement calculer des probabilités de choix.

On parle aussi de **modèles d'utilités aléatoires**.

Les alternatives (les options de choix) forment l'ensemble de choix \mathcal{K} .

Propriétés :

- les alternatives doivent être mutuellement exclusives (on ne peut en choisir qu'une) ;
- l'ensemble doit être exhaustif (i.e. représenter l'ensemble des choix possibles) ;
- de cardinalité finie.

Index des alternatives : $1, \dots, K$.

Notons $\mathbf{U} = (U_1, \dots, U_K)$: vecteur d'utilités associées aux alternative $1, \dots, K$.

Comme énoncé précédemment, l'utilité de chaque alternative peut être exprimée comme une fonction des caractéristiques observées de l'alternatives et des caractéristiques observées du preneur de décision.

Soit β le vecteur de variables reprenant ces caractéristiques et attributs, de sorte que

$$U_k = U_k(\beta).$$

Pour tenir compte des attributs et caractéristiques non-observées, nous décomposons les utilités comme suit :

$$U_k(\beta) = V_k(\beta) + \epsilon_k(\beta), \forall k \in \mathcal{K}.$$

Plus précisément :

- $V_k(\beta)$: partie déterministe,
- $\epsilon_k(\beta)$: partie aléatoire.

Remarque : le terme aléatoire est toujours supposé agir sous forme additive.

Nous supposons de plus qu'en moyenne, la partie déterministe représente adéquatement l'utilité, autrement dit,

$$E[\epsilon_k(\beta)] = 0,$$

ou encore,

$$E[U_k(\beta)] = V_k(\beta).$$

Dans ce contexte, nous utiliserons parfois la terminologie suivante :

- $U_k(\beta)$: utilité perçue,
- $V_k(\beta)$: utilité mesurée.

Fonction de choix (suite)

Supposons pour l'heure que la forme précise de U_k est connue, et est seulement fonction de β .

La probabilité de choisir k est noté P_k est aussi fonction de β , et est appelée *fonction de choix*.

P_k peut aussi être vue comme la proportion des individus, au sein d'une large population, qui sont caractérisés par β et qui choisissent k .

De plus, nous avons, $\forall I \in \mathcal{K}$,

$$P_k[\beta] = P[U_k(\beta) \geq U_l(\beta), \forall l \in \mathcal{K}].$$

Une fois la distribution du terme d'erreur ϵ_k spécifiée, la distribution des utilités peut être déterminée.

Considérons, par exemple, une situation où seulement deux alternatives sont disponibles, avec

$$U_1 = 3 + \xi,$$

$$U_2 = 2.$$

Dès lors,

$$P_1 = P[U_1 \geq U_2] = P[3 + \xi \geq 2] = P[\xi \geq -1].$$

Supposons à présent que

$$\xi \sim [-2, 2].$$

Dès lors, $P_1 = 0.75$, $P_2 = 0.25$.

Le probabilité de choix dépend par conséquent des attributs β , ainsi que de la distribution de la partie aléatoire. Le choix de la distribution permettra de caractériser plusieurs grandes familles de modèles.

Types usuels de modèles de choix discrets

- probit
- logit
- generalized extreme value (GEV)
- mixed logit

Types usuels de modèles de choix discrets

L'exemple précédent met aussi en avant un principe fondamental dans ce qui suivra.

Les utilités sont définies à une constante près. Ce qui importe, ce sont leurs différences, non leurs valeurs exactes. Les paramètres qui peuvent être estimés sont ceux qui capturent les différences entre alternatives.

Supposons à présent que les ϵ_k soient continues ($k = 1, \dots, K$), et notons

$$\epsilon = (\epsilon_1, \epsilon_2, \dots, \epsilon_K).$$

Par conséquent, ϵ possède une densité $f(\epsilon)$.

Différence des termes d'erreur

La probabilité de choix devient

$$\begin{aligned}P_k[\beta] &= P_\epsilon[U_k(\beta) \geq U_l(\beta), \forall l \in \mathcal{K}] \\&= P_\epsilon[V_k(\beta) + \epsilon_k \geq V_l(\beta) + \epsilon_l, \forall l \in \mathcal{K}] \\&= P_\epsilon[\epsilon_l - \epsilon_k \leq V_k(\beta) - V_l(\beta), \forall l \in \mathcal{K}]\end{aligned}$$

Nous devons donc observer toutes les différences entre paires d'alternatives, l'alternative choisie étant fixée. Dans le cas où $\#\mathcal{K} = 2$, la probabilité de choix est la fonction de répartition de la variable définie par la différence des termes d'erreur.

Différence des termes d'erreur (suite)

Réécriture :

$$P_k[\beta] = \int (I(\epsilon_l - \epsilon_k \geq V_k(\beta) - V_l(\beta) + \epsilon_l, \forall l \in \mathcal{K}) f(\epsilon) d\epsilon,$$

où I est la fonction indicatrice.

C'est une intégrale multidimensionnelle. Comment la calculer ?

Cela dépendra du choix des distributions :

- solution analytique,
- simulation.