

L'analyse conjointe

L'analyse conjointe définit un ensemble de techniques d'évaluation de certains biens, notamment les biens publics, les produits nouveaux et les services. L'AC cherche à quantifier l'appréciation d'un bien par le consommateur en le caractérisant par ses attributs et en étudiant les arbitrages que les individus opèrent entre différents *niveaux* de ces derniers. Il est possible d'utiliser l'analyse conjointe pour obtenir la disponibilité à payer (DAP) pour un nouveau produit en mettant un place un mécanisme *ad hoc* de révélation du choix. Un postulat essentiel de l'AC c'est que la valeur attribuée à chaque produit est une composition des valeurs associées à chaque niveau des attributs (caractéristiques) du produit même. La valeur se définit par une agrégation de ces attributs en une utilité subjective. L'analyse conjointe est également utilisée pour le développement d'un produit, la formation du prix d'un bien, le positionnement stratégique d'une entreprise, l'allocation des ressources publiques et le marketing industriel.

Introduction

Le terme analyse conjointe a été introduit pour la première fois par Green et Srinivasan (1978) pour regrouper techniques et modèles mettant en évidence la transformation de réponses subjectives à des stimuli (objectif) en des paramètres estimés. L'analyse conjointe est un développement naturel des mesures conjointes déjà appliquées en psychologie mathématique dans les années 60¹. Dans ces applications, la théorie de la mesure conjointe investiguait les relations fonctionnelles entre les stimuli provenant de multiples attributs caractérisant un bien et la valeur subjective (telle que perçue par l'individu) du bien-même. Une des hypothèses fondamentales est donc que le choix du consommateur entre des produits alternatifs se base sur la *valeur subjective* assignée à chacune de ces alternatives. Il est, par exemple, rationnel de croire que le consommateur décide d'acheter la voiture qui a la plus haute valeur subjective (utilité) entre tous les modèles de voitures disponibles sur le marché, et cela même s'il est impossible pour le modélisateur de prendre en compte tous les facteurs que l'individu gère dans la formation du choix. La compréhension du côté subjectif et comportemental, par un mécanisme de révélation des préférences, constitue le domaine de l'AC.

Plus formellement, l'AC utilise les formes fonctionnelles proposées par la théorie de la mesure conjointe pour relier l'utilité du produit à la valeur de l'attribut du produit. La procédure d'agrégation de l'utilité associée à chaque attribut, qui est au cœur de l'intérêt du chercheur, est connue comme **modèle de préférence**. Dans un tel contexte l'économétrie trouve son application avec de modèles d'utilité aléatoire (Random Utility Models, RUM) avec données discrètes. En assumant un modèle de préférence, le but de l'analyse conjointe est d'estimer l'importance de chaque combinaison attribut-niveau, en commençant des mesures de préférence explicites pour chaque produit alternatif.

¹ Luce et Tukey, (1964) "Simultaneous Conjoint Measurements: A New Type of Fundamental Measurement", Journal of Mathematical Psychology, 1, 1-27.

1 - L'AC et l'évaluation contingente

Le besoin de donner un prix à des biens non marchands n'est pas nouveau ; en 1952 Ciriacy-Wantrup avait déjà songé à des interviews pour estimer la valeur d'une réserve. Toutefois, il a fallu attendre Davis (1963) pour la mise en œuvre d'un questionnaire structuré sous forme d'enchères pour l'obtention de la valeur maximale des bénéficiaires récréatifs dont profitent les visiteurs d'une réserve naturelle. Le problème est double : d'un côté les individus peuvent déclarer plus qu'ils ne seraient effectivement prêts à payer, puisqu'ils ne sont pas vraiment face à une dépense ; d'autre part les gens peuvent très bien déclarer un montant moindre dans un comportement de type « resquilleur ».

Il est à remarquer que la méthode d'évaluation contingente a subi une évolution majeure dans les derniers dix ans en tant qu'instrument de quantification des dommages environnementaux suite à des accidents majeurs, tel le désastre du Exxon Valdez en Alaska. C'est à cette époque que la National Oceanic and Atmospheric Administration (NOAA) a commissionné à un panel d'économistes de mettre au point un « standard » de la méthode².

La méthode d'évaluation contingente se distingue clairement de l'AC par le fait que dans la première la valeur du bien analysé est demandée directement, tandis qu'en AC la procédure de révélation du choix est plus sophistiquée et en liaison directe avec les attributs du bien en question ; ce qui permet de corriger le danger d'une sous-estimation de la valeur du bien.

Les principales applications sont :

- Biens et « problèmes » environnementaux
- Nouveaux produits
- Evaluation de services et politiques

2 - Etapes de l'AC

Il est possible de schématiser les cinq étapes fondamentale de l'AC ainsi :

1^{re} Étape : Identification des caractéristiques

Les caractéristiques peuvent être identifiées par plusieurs méthodes. S'il s'agit d'une opinion concernant une politique, les caractéristiques seront prédéterminées par revue de la littérature, discussions de groupe collectives et pré enquêtes.

2^e Étape: Attribution des niveaux aux caractéristiques

² Arrow K. et al. (1993), « Report of the NOAA Panel on Contingent Evaluation », Federal Register, No. 58, pp 4601-4614.

Les niveaux des caractéristiques peuvent être cardinaux (par ex. « temps d'attente », où deux semaines est le double d'une), ordinaux (par ex. « douleur forte » est pire que « douleur modérée » mais on ne sait pas de combien), ou catégorique (lorsqu'il n'y a pas d'ordre du tout). Les niveaux doivent être plausibles et opérationnels de sorte à que les répondants prennent le questionnaire sérieusement.

3^e Étape : Choix des scénarios

Les scénarios sont formulés de sorte à décrire toutes les configurations étant donné les caractéristiques et niveaux choisis. Le nombre de scénarios augmente avec le nombre de caractéristiques et niveaux. En pratique, tous les scénarios générés peuvent rarement être inclus dans le questionnaire, pour cela des designs expérimentaux sont utilisés afin de réduire le nombre à un niveau acceptable.

4^e Étape: Établissement des préférences

Les préférences pour les scénarios inclus dans le questionnaire peuvent être révélées en utilisant une des trois méthodes :

- **Ordre** : les individus doivent lister les scénarios en ordre de préférence. La méthode a été employé surtout en marketing, mais pas, par exemple, dans le domaine de la santé;
- **Appréciation** : la méthode requiert que les répondants assignent une note, par ex. de 1 à 5, à chaque scénario ;
- **Choix discret** : les individus sont confrontés avec une série d'alternatives, et pour chacune il leur est demandé de choisir la préféré. Les réponses possibles incluent déclarer « A ou B est préféré », ou noter sur une échelle de 1 à 5, où 1 signifie « préfère fortement A » et 5 « préfère fortement B ».

5^e Étape : Analyse des données

Différentes techniques sont utilisées pour l'analyse des réponses. La méthode la plus appropriée est déterminée par le type de données récolté ; on distingue entre :

- Procédures métriques telles la régression multiple MCO, quand la variable dépendante est échelonnée par intervalle),
- Algorithmes non métriques tels MONANOVA³, LINMAP⁴ ou PREFMAP⁵, lorsque la variable dépendante est de type ordonné ;

³ Kruskal, J. B. (1965) Analysis of Factorial Experiments by Estimating Monotone Transformations of the Data. J. of the Royal Statistical Society, Series B, Vol. 27, pp. 251-263.

⁴ Srinivasan, V. and Shocker, A. D. (September 1973) Linear Programming Techniques for Multidimensional Analysis of Preference, Psychometrika. pp. 337-369.

⁵ Carrol (1972)

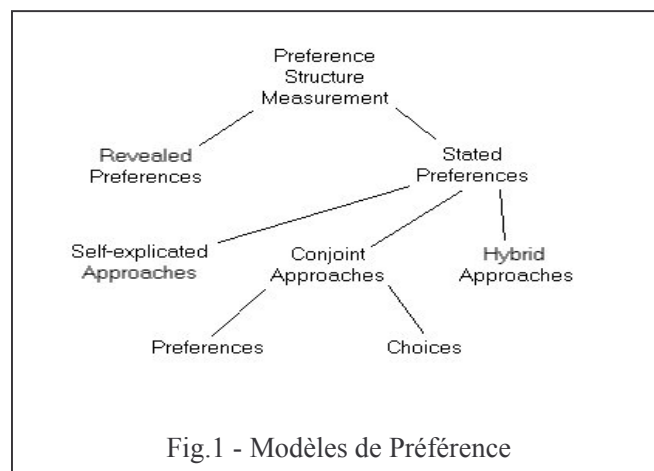
- Modèles logit mixte (MMNL) ou multinomiale (MNL), dans les cas des variables qualitatives pures.

3 - Modèles conjoints

Green et Wind⁶ ont été les premiers à avoir adressé le problème du choix entre alternatives avec multiples attributs dans la recherche en marketing. En effet, le marketing constitue un contexte privilégié de choix multiple dans lequel les modèles conjoints fournissent un outil puissant pour la mesure de la structure des préférences.

Le postulat essentiel de l'analyse conjointe c'est que la valeur attribuée à chaque produit est une composition des valeurs associées à chaque niveau des attributs (caractéristiques) du produit-même ; par exemple, vitesse maximale, consommation d'essence, taille du moteur et prix sont les attributs que le consommateur prend en considération dans l'évaluation d'une voiture avant l'achat. La valeur se définit ainsi par une agrégation de ces attributs en une utilité subjective⁷. L'information est généralement recueillie par moyen d'un questionnaire où on demande à chaque individu échantillonné de prendre en compte plusieurs attributs et produits simultanément.

La classification suivante des modèles pour la mesure de la structure des préférences a été proposée par Green et Srinivasan⁸:



Les modèles auto-explicatifs et hybrides sont alternatifs à la méthode conjointe. Dans l'approche auto-explicative le consommateur évalue chaque niveau de chaque attribut séquentiellement et indépendamment des autres combinaisons attribut-niveau.

⁶ Green P. E. and Wind Y. (1973) "Multiattribute Decisions in Marketing: A Measurement approach." Hinsdale, IL: The Dryden Press, 396pp.

⁷ Le prix est – évidemment - une « désutilité ».

⁸ Conjoint analysis in Marketing: New developments with implications for research in practice. J. of Marketing Oct. 1990.

L'approche hybride comporte deux tâches : une qui requiert au consommateur d'évaluer les attributs indépendamment les uns des autres et une autre où il doit produire une sorte d'exercice conjoint.

L'utilisation de critères de comparaison comme la qualité d'ajustement (goodness-of-fit) et finesse prédictive (accuracy) dans la recherche semble juger les modèles conjoints meilleurs que les auto-explicatifs, tandis que des différences marginales distinguent les approches conjointe et hybrides⁹.

Il existe deux approches distinctes pour la structure des préférences du consommateur¹⁰ :

- (a) l'approche basée sur les *préférences*, où le consommateur doit « noter » ou ordonner chaque produit alternatif,
- (b) l'approche basée sur les *choix*, où le consommateur doit choisir un produit entre plusieurs alternatives.

Dans l'approximation du choix réel du consommateur des raisons théoriques ont été apportées par Louviere et Woodworth¹¹ pour préférer l'utilisation des modèles basés sur les choix, plutôt que les préférences, : il n'existe pas de théorie formelle reliant les jugements au choix : il n'y a pas de logique explicite pour démontrer comment le consommateur lie ses préférences aux choix effectivement opérés : le processus de choix est probabiliste en nature et il ne peut pas être simulé par des équations conjointes déterministes. Donc, en utilisant des modèles conjoints basés sur les préférences il n'est pas possible d'estimer à quel degré la **présence** (ou l'absence) d'alternatives dans le set de choix influence le *comportement* de choix. La remarque est particulièrement importante en cas d'études d'analyse de produits concurrents. L'effet du nombre des niveaux (*number of levels effect*) dans les attributs est aussi un problème sérieux dans l'analyse conjointe¹².

⁹ Neslin, S. A. (1981) "Linking Product Features to Perceptions: Self-Stated Versus Statistically Revealed Importance Weights." J. of Marketing Research. Vol. 18, pp. 80-86 ; Cattin, P., Hermet, G. and Pioche, A. (1982) "Alternative Hybrid Models for Conjoint Analysis: Some Empirical Results." in Analytic Approaches to Product and Marketing Planning: The Second Conference, Edited by Srivastava, R. K. and Shocker, A. D. Cambridge, MA: Marketing Science Institute, October, Report No. 82-109, pp.142-151 ; Akaah, I. P. and Korgaokar, P. K. (May, 1983) "An Empirical Comparison of the Predictive Validity of Self-Explicated, Huber-Hybrid, Traditional Conjoint and Hybrid Conjoint Models." J. of Marketing Research. Vol.20, pp. 187-197.

¹⁰ Vriens M. (1995) "Conjoint Analysis in Marketing, Developments in Stimulus Representation and Segmentation Methods", Ridderprint, Ridderkerk, the Netherlands.

¹¹ Louviere J. J. and Woodworth G. (1983) "Design and Analysis of Simulated Consumer Choice or Allocation Experiments: An Approach Based on Aggregate Data". J. of Marketing Research. Vol. 20, pp. 350-367

¹² Voir "The number of levels effect – A proposed solution" par Dick Mc Cullough, 1999 Sawtooth Software Conference Proceedings.

En plus, seulement un nombre limité d'études de comparaison entre les approches conjointes basées sur les préférences et le choix ont été menés jusqu'à maintenant (Elrod, Louviere et Davey, 1992; Huber et. *al.*, 1992; Louviere et Gaeth, 1988; Oliphant et al., 1992) et, en plus, la plupart de ces études comparent des modèles de choix au niveau agrégé avec des modèles basés sur les appréciations au niveau individuel, en introduisant un biais naturel dû à l'ignorance de segments existant sur le marché. L'article de Moore, Louviere et Gray-Lee (1996) constitue une exception en comparant également différents modèles conjoints de segmentation.

3.1 - Modèles de Préférence

Dans le contexte des approches conjointes, trois modèles alternatifs sont proposés pour estimer les préférences¹³ :

1 - Le **modèle vecteur** assume la forme fonctionnelle linéaire - potentiellement restrictive – entre la préférence S_j pour le j^{e} profil et le niveau du p^{e} attribut pour le j^{e} profil y_{jp} ,

$$S_j = \sum_{p=1}^t W_p y_{jp}$$

où :

w_p : poids de l'individu pour le t attributs et y_{jp} est une variable continue (*e.g.* prix ou temps d'un trajet).

2 - Le **modèle du point-idéal** fait l'hypothèse que la préférence S_j inversement proportionnelle à la distance d_j^2 entre le point réel du j^{e} profil y_{jp} et son point idéal x_p (*i.e.* sa valeur idéale).

$$d_j^2 = \sum_{p=1}^t W_p (y_{jp} - x_p)^2$$

La relation de ce modèle implique que si d_j^2 devient plus petit S_j augmente et le profil évalué est plus proche du point idéal.

3 - Le **modèle part-worth**, basé sur la forme fonctionnelle suivante :

$$S_j = \sum_{p=1}^t f_p(y_{jp})$$

où f_p permet une estimation des coefficients pour chaque niveau du p^{e} attribut du j^{e} profil.

¹³Green P.E. and Srinivasan V. (1978), "Conjoint Analysis in Consumer Research: Issues and Outlook", *J. of Consumer Research*, 5 (September), 103-123.

Le modèle le moins restrictif est le modèle part-worth, il requiert typiquement l'estimation de $k \cdot t$ paramètres, où k est le nombre de niveaux dans le design expérimental et t est le nombre total d'attributs.

Le modèle vecteur requiert l'estimation de t paramètres $\{w_p\}$, alors que le modèle à point-ideal engage l'estimation $2 \cdot t$ paramètres, c'est-à-dire $\{w_p\}$ et $\{x_p\}$.

Le but étant l'estimation des préférences du consommateur, la sélection du modèle à utiliser devrait être basée sur la précision prédictive observée en utilisant chaque formulation dans le contexte expérimental spécifique.

3.2 - Modèles de Choix

Une autre catégorie de modèles de structure des préférences du consommateur basé sur les préférences déclarées est l'approche de choix conjoint, proposé pour la première fois par Louviere et Woodworth (1983)¹⁴.

L'hypothèse fondamentale est ici que le consommateur fait face à un set d'alternatives sur lesquelles il choisit ; les alternatives consistent en différentes combinaisons de niveaux sur un set de multiples attributs. La forme fonctionnelle utilisée le plus souvent est le modèle logit multinomial ou ses variantes.

La formule est :

$$P(a|A) = \frac{e^{V_a}}{\sum_{j \in A} e^{V_j}}$$

où $P(a|A)$ est la probabilité conditionnelle que le profil a sera choisi, sachant que le choix se fait à partir du set de choix A . v_a et v_j sont des variables explicatives dans les modèles part-worth. Le modèle part-worth, proposé par Louviere et Woodworth (1983) postule que la valeur V_j associée à chaque profil j dans le set de choix J est donnée par :

$$V_j = \beta_{0j} + \sum_k \beta_{kj} X_k$$

où

β_{0j} : valeur associée à l'alternative j indépendamment des attributs du profil et elle peut être vue comme l'évaluation des attributs non explicites ou latents;

β_{kj} : importance relative de l'attribut k tel qu'exprimé au niveau x_k à la valeur totale du profil v_j .

Une implication fondamentale du modèle MNL est l'indépendance des alternatives non pertinentes (l'hypothèse IIA) : la probabilité de choisir une alternative plutôt qu'une autre n'est pas influencée par les autres alternatives dans le set de choix. Il s'agit d'un facteur

¹⁴ Louviere J. J. and Woodworth G. (1983) "Design and Analysis of Simulated Consumer Choice or Allocation Experiments: An Approach Based on Aggregate Data". J. of Marketing Research. Vol. 20, pp. 350-367.

limitatif du modèle MNL qui peut être violé assez facilement. Des modèles alternatifs qui évitent cette limitation sont le logit emboîté ou le probit multinomial. Ces modèles demandent toutefois des procédures d'estimation plus complexes¹⁵.

Si, par contre, on suppose le MNL valable, on peut également rendre aléatoire (randomize) la séquence des sets de choix selon différents ordres afin de réduire les biais possibles par la violation de l'hypothèse IIA. Dans un tel cas la structure du questionnaire est conçue pour minimiser les effets de biais (tels *l'ordre des taches* et *la méthode de partage de variance*) par usage de séquences de profils générés aléatoirement et en alternant l'ordre des taches (s'il y en a plusieurs, par ex. appréciation et choix).

4 - Procédures d'estimation

La procédure d'estimation des paramètres du modèle dépend des hypothèses faites sur l'échelle et la distribution de la variable dépendante. Les procédures d'estimation suivantes sont applicables :

1 - Procédures métriques telles la régression multiple MCO, (la variable dépendante doit être - au moins - échelonnée par intervalle),

2 - Algorithmes non métriques tels MONANOVA¹⁶, LINMAP¹⁷ ou PREFMAP¹⁸ (Carroll, 1972), lorsque la variable dépendante est - au moins - sur une échelle ordinale,

3 - Modèles logit mixte (MMNL) ou multinomiale (MNL), lorsque la variable dépendante résulte des résultats du choix dans un set prédéfini.

4.1 - Algorithmes métriques vs. Non-métriques

Parmi les procédures énoncées dessus, les algorithmes métriques ont l'avantage de fournir des erreurs standard pour les paramètres estimés, en permettant des tests sur la significativité de chaque coefficient estimé. En plus, certaines procédures non métriques ne permettent pas à l'analyste d'imposer des contraintes spécifiques au modèle, par exemple, avec l'algorithme MONANOVA il est possible d'estimer seulement les modèles part-worth, i.e. il n'est pas possible d'imposer des contraintes **entre-attributs**. Il faut ajouter que dans les procédures non métriques il est seulement possible de spécifier des effets d'interaction *crossover*, alors que les algorithmes métriques permettent aussi l'inclusion des effets d'interaction non-crossover (Vriens, 1995).

¹⁵ Voir : Ben-Akiva M., McFadden D., Abe M., Böckenholt U., Bolduc D., Gopinath D., Morikawa T., Ramaswamy V., Rao V., Revelt D. and Steinberg D. (1997) "Modeling Methods for Discrete Choice Analysis" Marketing Letters. Vol 8, No 3, pp. 273-286.

¹⁶ Kruskal, J. B. (1965) Analysis of Factorial Experiments by Estimating Monotone Transformations of the Data. J. of the Royal Statistical Society, Series B, Vol. 27, pp. 251-263.

¹⁷ Srinivasan, V. and Shocker, A. D. (September 1973) Linear Programming Techniques for Multidimensional Analysis of Preference, Psychometrika. pp. 337-369.

¹⁸ Carroll (1972)

Autres différences existent : dans le cas de procédures métriques le critère du R^2 est typiquement utilisé dans un but d'optimisation et – avec certaines hypothèses vérifiées – il est possible de mener un F-test pour évaluer la significativité globale du modèle. Les algorithmes non métriques, par contre, ne supposent pas une distribution dans l'échantillon. Dans un tel cas, le test de la validité du modèle devient nettement plus subjectif et dépendant de la spécification des paramètres. Les méthodes utilisées pour maximiser le critère de significativité sont aussi également basés sur la minimisation d'une certaine distance entre les données observées et estimées. De toute manière ces méthodes sont généralement itératives et dans certains cas donnent des maxima locales.

La comparaison empirique des procédures d'estimation métriques et non métriques ont montré que même si l'utilisation des MCO données non métriques est théoriquement inappropriée, l'usage de procédures non métriques ne fournit pas de résultats significativement meilleurs (Vriens, 1995). Les algorithmes métriques affichent une performance proche des procédures non métriques notamment quand le ratio entre la taille de l'échantillon et le nombre de paramètres est petit (Green and Srinivasan, 1978). Cela est, bien sur, le cas pour l'estimation individuelle, mais il en est ainsi également pour les estimations au niveau agrégé ou segmenté.

4.2 - Estimation métrique

Plusieurs procédures peuvent être employées pour ajuster des modèles d'analyse conjointe avec ou sans segments. Vriens (1995) mentionne une étude où l'on teste 9 procédures alternatives selon les critères de qualité d'ajustement et finesse prédictive. Le résultat de l'étude le plus important révèle que la proportion de choix correctes est relativement constante, indépendamment de la procédure utilisée.

L'estimation des coefficients des segments des niveaux applique l'approche en deux étapes (Green et Krieger, 1991) : dans la 1^{re} étape on estime les coefficients individuels de régression ; lors de la 2^e, les individus sont agrégés dans des grappes construites selon la similarité des coefficients de régression estimés.

La procédure de clusterisation se base sur l'algorithme à k-moyennes (Hartigan, 1975 and MacQueen, 1967). Le principe c'est d'additionner les observation simples aux groups de manière itérative en optimisant un critère d'ajustement – typiquement le carrée de la distance entre les observations et le centre des clusters. Une autre procédure de segmentation métrique est basée sur l'optimisation du même critère mais, cette fois, pour l'estimation des paramètres de la fonction conjointe et les variables des segments. La procédure requiert toutefois un algorithme plus complexe.

4.4 - Estimation non métrique

Un cas spécifique d'estimation par procédure non métrique est donné par le maximum de vraisemblance, qui fournit des maxima globaux pour les paramètres du MNL. L'avantage principale de l'utilisation des estimateurs du maximum de vraisemblance, plutôt que les

MCO, c'est que même en cas de non respect des conditions de Gauss-Markov – comme c'est le cas pour variables dépendants binomiales ou multinomiales – la méthode fournit toujours des estimateurs qui sont asymptotiquement efficaces et consistants, auquel on peut appliquer des tests de signification.

Comme dans le cas d'autres modèles non métriques, le maximum de vraisemblance est une procédure d'estimation itérative lorsque des solutions en forme fermée – tels les MCO avec erreur normale iid – n'existent pas.

Un problème potentiel des estimateurs du maximum de vraisemblance appliqués au modèle MNL concerne la taille limitée des échantillons utilisés dans la recherche du consommateur ; ce qui pose des doutes quant aux propriétés asymptotiques des estimateurs ainsi que sur les tests d'hypothèses sur les paramètres. Malgré cela les modèles MNL et MMNL sont de plus en plus utilisés et ils affichent une bonne validité prédictive

Un développement récent dans l'utilisation du MNL avec des données de choix c'est la possibilité d'estimer les utilités des individus pour chaque segment de marché (voir e.g. DeSarbo, Ramaswamy and Cohen, 1995; Kaciak and Louviere, 1990; Louviere, 1994).

Louviere (1994) suggère l'emploi de mesures de *relatedness* appropriées aux données afin de comparer les choix des répondants, comme – par exemple - le nombre d'occurrences pour le même profil à partir des profils alternatifs. Un autre critère pour grouper les répondants en des catégories uniformes c'est l'application de l'analyse à correspondance multiple sur les données de choix pour réduire l'espace à peu de groupes auxquels on alloue les répondants. Dans les deux cas il faut employer des algorithmes de clusterisation et estimer un MNL séparément pour chaque segment.

Une technique de segmentation c'est le MNL avec classes latentes (DeSarbo, Ramaswamy and Cohen, 1995). La technique est attrayante, car elle fait recours à l'estimation simultanée des paramètres du modèle de choix et des indicateurs du segment; elle requiert toutefois la spécification d'une fonction de vraisemblance complexe qui est maximisée par des procédures numériques itératives.

5 – Exemples d'estimations

5.1 - Données d'appréciation

La variable dépendante pour des données d'appréciation est typiquement échelonnée par intervalle (interval-scaled) ; il s'agit de valeurs continues entre 1 et 10 (ou 1 et 100) pour laquelle il est possible l'emploi de moindres carrés simultanés.

Ayant pour but l'évaluation d'une boisson nous pouvons le caractériser par coût, goût et douceur ; un profil pourrait être :

Coût	4 CHF
Goût	Orange
Douceur	pas de sucre

un autre c'est :

Coût	3 CHF
Goût	Cola
Douceur	Sucre

Le chercheur peut s'intéresser à trois niveaux de prix – par ex. 3, 4, 5 CHF ; trois goûts (orange, cola et citron) et deux type de douceur (sucre/pas de sucre), pour aboutir à 18 profils différents. Il est dès lors évident qu'en augmentant attributs et niveaux le nombre de profils augmente très rapidement, ce qui amène des problèmes dans l'estimation. Pour réduire le nombre de profils il y a plusieurs méthodes, la principale consiste à « orthogonaliser » les profils.

Considérons le modèle linéaire :

$$Y_j = \beta_0 + \beta_1 X_{j1} + \beta_2 X_{j2} + \beta_3 X_{j3} + \varepsilon_j$$

où X_{ji} sont des attributs avec deux niveaux chacun.

Sans perte de généralité il est possible de coder les niveaux par -1 et +1 (niveau bas et haut), en aboutissant à huit profils :

Profil	X_1	X_2	X_3
1	-1	-1	-1
2	-1	-1	1
3	-1	1	-1
4	-1	1	1
5	1	-1	-1
6	1	-1	1
7	1	1	-1
8	1	1	1

Tableau 1- Plein profil

Si les profils 1 et 4 sont sélectionnés, l'effet de X_1 ne pourra pas être estimé. Pour obtenir un set orthogonal nous pouvons écrire toutes les combinaisons des deux premiers attributs et additionner les niveaux du troisième de sorte à avoir $X_{j1} * X_{j2} * X_{j3} = 1$. Le profil résultant est :

Profil	X ₁	X ₂	X ₃
1	-1	-1	1
2	-1	1	-1
3	1	-1	-1
4	1	1	1

Tableau 2 – Profil orthogonal

La propriété d'orthogonalité implique que les attributs sont non corrélés dans le sous-set de profils, formellement, en ayant $X_j = (1 \ X_{j1} \ X_{j2} \ X_{j3})^T$, alors :

$$\sum_{j=1}^4 X_j X_j^T = \text{diag} \left\{ \sum_{j=1}^4 X_{jk}^2 \right\}$$

Cela implique que les estimateurs MC des coefficients $\beta_0 \ \beta_1 \ \beta_2$ ne sont pas corrélés et la matrice des variances-covariances c'est :

$$\text{Cov}(\beta_{MC}) = \sigma^2 \left(n \sum_{j=1}^4 X_j X_j^T \right)^{-1}$$

avec n les nombre d'individus et σ^2 la variance de l'erreur du modèle linéaire.

Les profils orthogonaux sont très populaires, mais puisque les données sont souvent de type qualitatifs, les modèles linéaires résultent inappropriés.

Supposons d'avoir une variable dépendante binaire : $Y_j \in \{0, 1\}$, nous pouvons utiliser un modèle logit :

$$P(Y_j=1) = \frac{\exp\{X_j^T \beta\}}{1 + \exp\{X_j^T \beta\}}$$

la vraisemblance pour la i^{ème} observation c'est :

$$\log L(\beta)_i = \sum_{j=1}^4 X_j^T \beta - \log(1 + \exp\{X_j^T \beta\})$$

et la matrice des variances-covariances c'est :

$$\text{VarAsy}(\beta_{MV}) = \left(\sum_{j=1}^4 \frac{\exp\{X_j \beta\}}{1 + \exp\{X_j \beta\}} \right)^2 X_j X_j^T$$

Le profil spécifié jusqu'ici exclue des effets d'interaction, mais cette hypothèse peut s'avérer fausse en cas d'effets croisés. Il est possible de créer un profil avec des effets d'interaction : l'effet principal de X₁ est accompagné d l'effet joint de X₂ X₃ :

	Effets principaux			Effets croisés			
Profil	X ₁	X ₂	X ₃	X ₁ X ₂	X ₁ X ₃	X ₂ X ₃	X ₁ X ₂ X ₃
1	-1	-1	1	1	-1	-1	1
2	-1	1	-1	-1	1	-1	1
3	1	-1	-1	-1	-1	1	1
4	1	1	1	1	1	1	1

Tableau 3 – Profil avec effets d'interaction

5.2 – Mesure des préférences

En ayant $X_j = (1 \ X_{j1} \ X_{j2} \ X_{j3})^T$, le $j^{\text{ème}}$ profil, le $i^{\text{ème}}$ individu associe une préférence à ce profil selon une fonction $U_{ij} = f_i(X_j)$. La préférence U est une variable latente inconnue. Dans le cas d'expérience avec données d'appréciation une échelle d'appréciation est établie (par ex. de 1 à 7) sur laquelle l'interviewé note chaque profil. En cas de données ordonnées l'individu range tous les profils du plus au moins préféré ; s'il y a un grand nombre de profils le rangement des profils peut se faire en deux étapes : il peut par ex. prendre les cinq premiers et les ordonner et puis traiter les autres.

5.3 - Estimation du modèle de préférence

5.3.1 – Données d'appréciation

Considérons pour le $i^{\text{ème}}$ individu U_{ij} sa préférence pour le $j^{\text{ème}}$ profil. U_{ij} est mesuré sur une échelle de 1 (préférence basse) à M (préférence haute). Nous pouvons écrire :

$$Y_{ij} = y \quad \text{si} \quad c_{y-1} < U_{ij} < c_y \quad \text{pour } 1 < y < M$$

où,

$$c_0 < c_1 < \dots < c_M$$

et,

$$c_0 = -\infty$$

$$c_M = \infty$$

sont des valeurs « seuil ».

En assumant une fonction de préférence linéaire :

$$U_{ij} = X_j \beta + \varepsilon_{ij}$$

et, $g(\varepsilon)$, $G(\varepsilon)$ étant respectivement les fonction de densité et de répartition de ε , alors :

$$\begin{aligned}
P(Y_{ij} = y) &= P(c_{y-1} < U_{ij} < c_y) \\
&= P(c_{y-1} - X_j\beta < U_{ij} < c_y - X_j\beta) \\
&= G(c_y - X_j\beta) - G(c_{y-1} - X_j\beta)
\end{aligned}$$

Si ε_{ij} est normalement distribué (0, 1), nous obtenons un modèle probit ordonné, tandis que dans le cas où les ε_{ij} suivant une distribution de Valeur Extrême de type 1 nous obtenons le modèle logit. La fonction de vraisemblance pour le $i^{\text{ème}}$ individu c'est :

$$L_i(\beta) = \prod_{j=1}^J (G(c_{y_{ij}} - X_j\beta) - G(c_{y_{ij}-1} - X_j\beta))$$

L'estimateur du maximum de vraisemblance est la valeur qui maximise :

$$L(\beta) = \prod_{i=1}^n L_i(\beta)$$

avec l'hypothèse d'indépendance entre les observations.

Il est aussi possible d'introduire une composante aléatoire dans les paramètres pour inclure, par ex. des variations du aux goûts, des lors nous écrivons :

$$U_{ij} = X_j\beta + \alpha_i + \varepsilon_{ij}$$

où α représente la composante aléatoire de densité $h(\alpha ; \delta)$, où δ est la matrice var-cov. Dans un tel cas la vraisemblance pour une observation devient :

$$L_i(\beta, \delta) = \int_{-\infty}^{+\infty} \prod_{j=1}^J [G(c_{y_{ij}} - X_j\beta - \alpha) - G(c_{y_{ij}-1} - X_j\beta - \alpha)] h(\alpha, \delta) d\alpha$$

L'estimateur du MV $(\hat{\beta}, \hat{\delta})^T$ est défini comme la valeur qui maximise :

$$\log L(\beta, \delta) = \sum_{i=1}^n \log L_i(\beta, \delta)$$

5.3.2 – Données d'ordre

Avec des données ordonnées les préférences du $i^{\text{ème}}$ individu sont directement placés sur une échelle ordinale. Notons j_m le nombre de profil placé en rang m : le profil préféré est j_1 et le moins préféré est le j_J . Nous avons l'information :

$$U_{ij_1} > U_{ij_2} > \dots > U_{ij_J}$$

En prenant le modèle à effets aléatoires cela équivaut à :

$$X_{j1}\beta + \varepsilon_{ij1} > X_{j2}\beta + \varepsilon_{ij2} > \dots > X_{jJ}\beta + \varepsilon_{ijJ}$$

Il est à remarquer que la composante spécifique à l'individu α n'affecte pas le rang des profils.

Si les erreurs sont iid selon une loi de valeur extrême de type 1, alors :

$$P(j_1, \dots, j_J) = P(U_{j_1} > U_{j_2} > \dots > U_{j_J}) = \prod_{s=1}^J \frac{\exp\{X_{j_s}\beta\}}{\sum_{t=s}^J \exp\{X_{j_t}\beta\}}$$

La log-vraisemblance pour le $i^{\text{ème}}$ individu c est :

$$\log L_i(\beta) = \sum_{s=1}^J [X_{j_s}\beta] - \log \sum_{t=s}^J \exp\{X_{j_t}\beta\}$$

Sous des conditions appropriées, l'estimateur du MV est asymptotiquement normal et la matrice de variances covariances s'écrit :

$$\text{VarAsy}(\beta_{MV}) = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \left[E \left(\frac{\partial^2 \log L}{\partial \beta \partial \beta'} \right)^2 \right]^{-1}$$

5.4 – Evaluation du niveau des attributs

Il est possible d'évaluer le changement d'un niveau d'un attribut (par ex. le prix) en termes de variation compensatoire.

Prenons la fonction de préférence linéaire :

$$U_{ij} = \alpha_i + \sum_{s=1}^{k-1} X_{js}\beta_s + \gamma c_j + \varepsilon_{ij}$$

avec c_j le cost du $j^{\text{ème}}$ profil et γ son coefficient. Une mesure de la Disposition à payer pour un accroissement d'une unité du $s^{\text{ème}}$ attribut c est :

$$\text{DAP}_s = -\beta_s / \gamma$$

Une mesure consistante de la DAP s'obtient en remplaçant les coefficients inconnus avec les estimateurs.

La valeur économique du $j^{\text{ème}}$ profil c est :

$$\text{DAP}_{\text{profil } j} = -X_j\beta / \gamma$$

remarquons que la dernière expression est une somme pondérée des DAP marginales :

$$DAP_{profilj} = \sum_{s=1}^{k-1} X_{js} DAP_s$$

6 - Expériences avec données de choix

Les données d'appréciation et d'ordre dans la révélation des préférences des individus utilisent des méthodes qui ne correspondent pas nécessairement au comportement réel, qui parfois revient à faire un choix entre alternatives.

Dans une expérience de choix le chercheur construit des sets d'alternatives entre toutes les possibilités à analyser ; l'individu doit choisir la préférée. Si nous reprenons l'exemple de la boisson, nous avons :

Attribut	A	B	C
Coût	5	4	5
Goût	Orange	Cola	Citron
Douceur	Pas de sucre	Sucre	Sucre

Tableau 4 – Set de choix

L'individu doit simplement choisir un produit et pour plusieurs sets de choix.

6.1 – Modélisation avec données de choix

Nous nous plaçons dans le contexte de la théorie de maximisation de l'utilité stochastique. En ayant M profils dans le set de choix, le problème du consommateur consiste à :

$$\begin{aligned} \max Z, m U_i^* (Z, X_m) \\ \text{s.c. } ZP \leq Y_i - c_m \end{aligned}$$

avec U_i^* fonction d'utilité, Z vecteur des biens, P vecteur des prix, Y_i le revenu, X_m vecteur d'attributs du même profil et c_m coût du $m^{\text{ème}}$ profil (ici inclus dans X_m).

L'individu choisit *toujours* l'alternative qui engendre la plus grande utilité, même si *tous* les facteurs de choix ne sont pas observés par l'analyste.

Considérons une fonction d'utilité linéaire :

$$U_{im} = P\lambda + X_m\beta + \gamma(Y_i - c_m) + \alpha_i + \varepsilon_{im}$$

où λ et β sont des vecteurs de coefficients, γ est l'utilité marginale du revenu, α_i est une composante spécifique de l'individu et ε_{im} est un terme inconnu. Puisque $P\lambda$, γY_i et α sont constants entre les différentes alternatives, ils n'influencent pas le choix ; par conséquent la règle de décision s'exprime : l'individu choisit le profil m ssi :

$$X_m \beta - \gamma c_m + \varepsilon_{im} > X_j \beta - \gamma c_j + \varepsilon_{ij}, \forall j \in C; j \neq m$$

Si on fait l'hypothèse que le individus sont choisit de manière aléatoire, le terme spécifique à l'individu peut être traité comme une erreur aléatoire.

La probabilité qu'un individu choisisse l'alternative m s'écrit :

$$P(m:C, \beta, \gamma) = P(X_m \beta - \gamma c_m + \varepsilon_{im} > X_j \beta - \gamma c_j + \varepsilon_{ij}; \forall j \in C; j \neq m)$$

en faisant l'hypothèse que les erreur ont une fonction de densité $g(\varepsilon)$, alors :

$$P(m:C, \beta, \gamma) = \int_{-\infty}^{\infty} g(u - X_m \beta - \gamma c_m) \prod_{j \neq m}^M G(u - X_j \beta - \gamma c_j) du$$

Si le erreurs sont iid selon une distribution à valeur extrême :

$$P(m:C, \beta, \gamma) = \frac{\exp\{X_m \beta - \gamma c_m\}}{\sum_{j \in C} \exp\{X_j \beta - \gamma c_j\}}$$

6.2 - Mesures de la disposition à payer

Il existe différentes mesures de DAP ; une est donnée par :

$$DAP_s = \beta_s / \gamma$$

tandis que pour le profil j nous avons

$$DAP_{\text{profil } j} = X_j \beta / \gamma$$

Ces mesures sont valables s'il n'existe pas de substituts du produit ; si ce n'est pas le cas l'choix des alternatives doit être considéré : posons les J alternatives dont l'utilité s'écrit :

$$U_j = V_j + \varepsilon_j$$

V étant la composante systématique de l'utilité, i.e. $V_j = X_j \beta$

La variable indicatrice $I(j) = 1$ si la $j^{\text{ème}}$ alternative est choisie ($U_j > U_m \forall m$) ; l'utilité de l'individu s'écrit :

$$U = \sum_{j=1}^J I(j) U_j = \sum_{j=1}^J I(j) (V_j + \varepsilon_j)$$

l'espérance de cette expression pour un individu choisit au hasard c'est :

$$E(U) = E \sum_{j=1}^J I(j) (V_j + \varepsilon_j)$$

qui, dans le cas d'une distribution des erreurs selon une loi de valeur extrême devient :

$$E(U) = \log \sum_{j=1}^J \exp \{V_j\}$$

La DAP moyenne pour le changement de la valeur des attributs de X_j à X_j' est mesuré par :

$$DAP = \frac{1}{\gamma} \left(\log \sum_{j=1}^J \exp \{V_j'\} - \log \sum_{j=1}^J \exp \{V_j\} \right)$$

Cette expression est la même que dans le cas de préférences révélées. Si $J = 1$,

$$\log \sum_{j=1}^J \exp \{V_j\} = V_1 \text{ et } DAP = \frac{1}{\gamma} (X_1' - X_1) \beta$$

qui est la formule de calcul de la DAP dans l'analyse contingente.

- ● -

Annexe 1 - Préférences déclarées vs. Révélées

L'information de l'AC provient en général des préférences déclarées - *stated preferences*, ou SP – qui, en opposition aux préférences révélées - *revealed preferences*, RP – présentent plusieurs avantages :

- Les entreprises ont besoin de modéliser la demande de nouveaux produits avec nouveaux attributs pour lesquelles il ne peut pas y avoir de RP.
- Les variables explicatives obtenues par RP ont souvent peu de variabilité, ce qui peut poser des problèmes de colinéarité ; la colinéarité provient de deux sources liées :
 - 1) La **technologie du marché**, qui amène de la corrélation entre les attributs des produits et limite la variabilité des attributs-même.

- 2) Les **forces de marché**, qui encouragent les producteurs à offrir des produits avec attributs similaires, ainsi qu'à des corrélations négatives entre attributs pour une frontière d'efficacité donnée.
- L'évolution du marché introduit des nouveaux attributs et caractéristiques (qu'on cherche à analyser).
 - Les données SP sont souvent moins chères à obtenir
 - Des données RP pour certains biens n'existent pas (biens publics).

En résumant, les données RP sont axées sur les préférences présentes ou d'un passé récent, en termes de marché et technologie, tandis que les RP permettent de simuler un changement technologique profond et d'en quantifier les effets dans la demande future. Plusieurs applications commerciales de l'analyse conjointe ont eu lieu depuis la fin des années 70 et un véritable boom de la technique est observé depuis la fin des années 80, en coïncidence avec l'introduction de différents logiciels (par ex. Sawtooth™) pour le design de l'étude, la collecte et l'analyse des données conjoints¹⁹.

Annexe 2 – Echelles de mesure

Au sens large, la mesure consiste à attribuer des entités mathématiques, généralement des nombres, à d'autres entités, tels des phénomènes empiriques. Le but du chercheur est de structurer l'attribution de sorte à ce que la relation fondamentale entre les nombres représente celle entre les entités empiriques. Cette *correspondance*, quand elle est solide, permet d'utiliser les propriétés des mathématiques comme un modèle descriptif du phénomène empirique.

Echelle Ordinale

L'échelle ordinale est obtenue lorsque les relations entre les objets sont du type d'*ordre*, tels attitudes, préférences, croyances, perceptions. Typiquement on demande de classer n objets auquel on assigne les entiers 1,2,3... Il est possible d'avoir des intervalles différents entre les objets et avoir un set du type : 1,4,7,10. Cette possibilité préserve la relation d'ordre seulement si la séquence est monotone croissante. Quand les nombres assignés aux objets représentent seulement la relation d'ordre, on parle de données en échelle ordinale.

¹⁹ Voir : Wittink D. R., Vriens M., and Burhenne W. (1994) "Commercial Use of Conjoint Analysis in Europe: Results and Critical Reflections". International J. of Research in Marketing, Vol. 11, pp. 41-52.

Echelle d'Intervalle

Les nombres sont assignés pour représenter non seulement l'ordre mais aussi la distance entre les objets on parle d'échelle d'intervalle. Typiquement la différence entre l'objet A et B nous dit combien plus A (ou B) a par rapport à B. Dans une échelle d'intervalle, des différences égales entre les valeurs doivent refléter des différences empiriques égales entre les objets.

Echelle de Ratio

Si les nombres sont attribués au niveau d'une échelle d'intervalle et il est aussi possible d'établir un point zéro unique ou une origine commune, l'échelle est de type ratio, telles : poids, hauteur, longueur. Dans le comportement du consommateur elles sont rares.

Transformations permises

Il est possible de regarder aux différences entre les trois type d'échelle en termes de l'unicité de certaines transformations. L'artifice est utile pour passer de l'attribut de l'objet à l'utilité. L'échelle ordinale permet une transformation monotone croissante (pas forcément linéaire). Les transformations des échelle d'intervalle doivent être linéaires, tandis que les échelles de ratio doivent être linéaires et proportionnelles.

Echelle métrique ordonnée

Dans une échelle métrique ordonnée il est possible de ranger tous les intervalles entre pairs. En ayant, par exemple, $\{A, B, C, D, E\}$ ordonnés le long d'une ligne continue, il est possible de ranger toutes les dix distances : $\overline{AB}, \overline{AC}, \dots, \overline{DE}$. L'échelle est ordinale, mais avec l'information supplémentaire de l'ordre des distances entre les couples des points.

Giancarlo Fiorito