Plans d'expériences et analyse conjointe: applications en marketing

Gilbert Saporta

PLANS D’EXPÉRIENCES ET ANALYSE CONJOINTE: APPLICATIONS EN MARKETING

Gilbert Saporta

L’analyse conjointe est tout d’abord une méthode d’analyse des données destinée à relier une variable à expliquer ordinale à plusieurs variables explicatives qualitatives à modalités ordonnées ou non. En ce sens l’analyse conjointe est une forme particulière de régression multiple avec transformation des variables.

Dans une optique plus générale l’analyse conjointe est une méthodologie d’étude complète comportant une phase de recueil des observations basée sur un choix de combinaisons de variables déterminée par application de la technique des plans d’expériences, une phase de traitement destinée à estimer les paramètres, une phase de simulation.

Le domaine privilégié d’application se situe en marketing et on parle alors de méthodes de « Trade Off ». Il s’agit d’expliquer les préférences de consommateurs pour des produits définis par des combinaisons d’attributs. Chaque consommateur enquêté note ou classe un ensemble de produits réels ou hypothétiques. L’analyse conjointe décompose alors les préférences selon un modèle d’utilité additive propre à chaque interviewé.

Le terme de « Trade Off » évoque un modèle compensatoire où une bonne note sur un attribut peut être compensée par une note moins bonne sur un autre attribut : le consommateur effectue son choix de produits en mettant en balance avantages et inconvénients .

1- LES DONNÉES ET LE MODÈLE

Partons de l’exemple suivant : un enquêté classe par ordre de préférence 12 produits décrits par 4 variables qualitatives A, B, C, D à 3, 3, 4 et 4 modalités respectivement :

n°					rang
1	A1	B1	C1	D1	7
2	A2	B1	C1	D2	8
3	A3	B2	C1	D3	10
4	A1	B1	C2	D2	5
5	A2	B2	C2	D3	1
6	A3	B3	C2	D4	11
7	A1	B1	C3	D3	4
8	A2	B3	C3	D4	2
9	A3	B1	C4	D4	12
10	A1	B1	C4	D1	6
11	A2	B1	C4	D1	3
12	A3	B2	C4	D2	9

Ces produits étaient des services possibles pour la vente d’appareils de télévision :

A modalités de retrait :

A1 appareil pris en rayon, paiement à la caisse
A2 appareil remis après paiement
A3 appareil remis à un autre comptoir après paiement

B Prestation

B1 vendu sans essai
B2 essai
B3 essai et démonstration

C durée de garantie

C1 6 mois
C2 1 an
C3 2 ans
C4 3 ans

D délais de réparation

D1 dans les 24 heures
D2 dans la semaine
D3 dans la quinzaine
D4 dans le mois

On fait l'hypothèse que le classement du consommateur résulte d'une addition de points, « les utilités », associés aux différentes modalités présentes dans les produits (ce qui revient à considérer qu'il arbitre d'une manière additive entre les qualités constituant les produits).

Si y_6 désigne le rang du produit n°6, on doit avoir approximativement :

$$f(y_6) = \alpha_3 + \beta_3 + \gamma_2 + \delta_4$$

où f est une fonction monotone croissante.

Selon les cas on prendra pour f l'identité, ce qui revient à estimer directement les rangs de classement comme si c'était une variable numérique, ou bien on cherchera la transformation monotone la meilleure pour expliquer le classement.

Il y a ici autant de coefficients à estimer que de modalités, soit dans notre exemple 14 coefficients (donc plus que produits à classer). Lorsque f est égale à l'identité, on aura reconnu un modèle linéaire (sans constante) associé à une analyse de variance où seuls les effets principaux (additifs) sont présents.

Dans le cas général ce modèle s'écrit f(**y**) = **Xb** + **e** où **e** est le terme d'erreur et **X** la matrice d'expérience qui vaut :

```
1 0 0  1 0 0  1 0 0 0  1 0 0 0
0 1 0  1 0 0  1 0 0 0  0 1 0 0
0 0 1  0 1 0  1 0 0 0  0 0 1 0
1 0 0  1 0 0  0 1 0 0  0 1 0 0
0 1 0  0 1 0  0 1 0 0  0 0 1 0
0 0 1  0 0 1  0 1 0 0  0 0 0 1
1 0 0  1 0 0  0 0 1 0  0 0 1 0
0 1 0  0 0 1  0 0 1 0  0 0 0 1
0 0 1  1 0 0  0 0 1 0  1 0 0 0
1 0 0  1 0 0  0 0 0 1  0 0 0 1
0 1 0  1 0 0  0 0 0 1  1 0 0 0
0 0 1  0 1 0  0 0 0 1  0 1 0 0
```

Cette matrice comporte autant de lignes que de produits, ou scénarios, classés et possède un nombre de colonnes égal à la somme des nombres de modalités des variables explicatives encore appelées facteurs. Les colonnes indiquent pour chaque produit les modalités prises sous forme logique (variables indicatrices).

Comme dans tous ces types de problèmes, il y a indétermination des coefficients car les colonnes de **X** sont liées par des relations linéaires : la somme des indicatrices des modalités de chaque facteur vaut toujours 1.

Ceci signifie que seules les différences entre coefficients ont une signification :

$\alpha_1 - \alpha_2$ est estimable mais pas α_1 et α_2 .

En pratique on imposera des contraintes aux coefficients ; les plus courantes étant soit d'avoir des coefficients de somme nulle $\alpha_1 + \alpha_2 + \alpha_3 = 0$, soit d'annuler un des coefficients, en général le premier ou le dernier ce qui revient à éliminer des colonnes de **X**.

Notons qu'il est facile de passer d'une solution où l'un des coefficients est nul à la solution à somme nulle : il suffit de retirer la moyenne, autrement dit de centrer les résultats. Dans l'exemple précédent il y a donc 14 - 4 = 10 coefficients estimables indépendamment les uns des autres pour 12 observations, ce qui laisse seulement deux degrés de liberté.

Remarquons que l'écriture précédente du modèle conduit à des utilités qui sont en fait des anti-utilités car les meilleurs classements sont associés à des valeurs faibles de y et donc à des coefficients faibles aussi. On procédera si l'on veut de vraies utilités à la transformation suivante des rangs : $y = m_y - (y - m_y)$ où m_y est le rang moyen.

Il est plus courant en pratique de demander de classer les produits selon un ordre complet plutôt que de les noter (sur 20 par exemple), ceci afin d'éviter les ex-aequo et pour forcer les enquêtés à comparer réellement les produits.

2- L'ESTIMATION DES UTILITÉS INDIVIDUELLES

Le modèle précédent est appliqué à chacun des enquêtés et fournira donc des utilités propres à chacun. Il ne s'agit pas d'un modèle général de préférences mais bien d'un modèle tentant d'analyser les préférences de chaque consommateur.

Techniquement il faudra donc réaliser autant de régressions, ordinaires ou monotones, qu'il y a d'enquêtés. Il convient donc d'être vigilant sur la représentativité de l'échantillon et d'éviter des simplifications abusives en calculant trop vite des utilités moyennes, comme on le verra plus loin.

Pour estimer les utilités on utilisera soit la méthode des moindres carrés ordinaires, soit une méthode de régression monotone.

2.1 Estimation par les moindres carrés ordinaires

On part ici de $\mathbf{y} = \mathbf{Xb} + \mathbf{e}$ où $\mathbf{y}$ est le vecteur des rangs de classements des scénarios et on estime $\mathbf{b}$ par $\mathbf{b}$ de telle sorte que :

$$\| \mathbf{y} - \mathbf{Xb} \|^2 \text{ soit minimal.}$$

Techniquement on centrera les rangs en retirant le rang moyen car le modèle est ici sans constante.

Pour les raisons déjà évoquées plus haut, ces estimations ne sont pas uniques, et on obtient une solution simple en éliminant des colonnes de **X.**

On prendra par exemple $\mathbf{X}_0$ égal à :

0 0	0 0	0 0 0	0 0 0
1 0	0 0	0 0 0	1 0 0
0 1	1 0	0 0 0	0 1 0
0 0	0 0	1 0 0	1 0 0
1 0	1 0	1 0 0	0 1 0
0 1	0 1	1 0 0	0 0 1
0 0	0 0	0 1 0	0 1 0
1 0	0 1	0 1 0	0 0 1
0 1	0 0	0 1 0	0 0 0
0 0	0 0	0 0 1	0 0 1
1 0	0 0	0 0 1	0 0 0
0 1	1 0	0 0 1	1 0 0

b est alors égal à $(\mathbf{X}'_0\,\mathbf{X}_0)^{-1}\,\mathbf{X}'_0\,\mathbf{y}$.

Dans l'exemple, on obtient les utilités suivantes après avoir retiré 6.5, valeur du rang moyen :

$\alpha 1 = 0$	$\alpha 2 = 0.1$	$\alpha 3 = 8$	
ß1 = 0	ß2 = -3.6	ß3 = -4.6	
$\gamma 1 = 0$	$\gamma 2 = -2$	$\gamma 3 = -2.5$	$\gamma 4 = -3.3$
$\delta 1 = 0$	$\delta 2 = 1.1$	$\delta 3 = -0.3$	$\delta 4 = 2.8$

On constatera que les scores obtenus par le modèle d'utilités additives sont dans le même ordre, sur cet exemple d'école, que les rangs de classement mais que ceux-ci ne sont pas reconstruits exactement le R^2 valant 0.9385.

2.2 Régression monotone

Considérer le rang de classement comme une variable numérique est discutable, il s'agit uniquement d'une information ordinale et tout système de notation qui conduit au même ordre est acceptable. Comme il s'agit de reconstituer l'ordre de préférences on peut faire subir aux rangs de classement une transformation monotone croissante. Le principe de l'analyse de variance monotone (programme MONANOVA de J.B.Kruskal) consiste à chercher la transformation monotone T(**y**) qui conduit à la meilleure régression multiple avec **X** (ou $\mathbf{X}_0$) comme matrice de variables explicatives.

Voyons rapidement comment on peut estimer la transformation monotone et les utilités à l'aide d'une méthode de moindres carrés alternés.

Tout d'abord examinons la structure des transformations monotones de **y.**

On supposera pour simplifier que les valeurs de **y** ont été réordonnées par valeurs croissantes. Une transformation monotone de **y** est donc de la forme :

$T(y_1) = a_1$

$T(y_2) = a_1 + a_2$
$T(y_3) = a_1 + a_2 + a_3$

...

$T(y_n) = a_1 + a_2 + ... + a_n$
où a_1 est de signe quelconque alors que les autres coefficients a_i sont positifs ou nuls .

On voit alors que T(**y**) s'exprime comme une combinaison linéaire de n variables à coefficients positifs (sauf un) ce qui donne à l'ensemble des transformations monotones de **y** la structure d'un cône polyédrique convexe.

La recherche de la minimisation de $\| T(\mathbf{y}) - \mathbf{Xb} \|^2$ est alors un problème d'analyse canonique entre ce cône et le sous-espace vectoriel W engendré par les colonnes de **X**. Ce problème peut se résoudre par une méthode de projections successives (cf M.Tenenhaus et A.Agha) :

Connaissant T, il s'agit d'une régression multiple ordinaire, projection sur Le sous-espace engendré par les colonnes de **X** ; **Xb** fixé on obtient T en projetant sur le cône, cette projection est unique et peut être obtenue soit par la méthode de Kruskal ou en appliquant l'idée suivante de M.Tenenhaus : projeter un vecteur **z** sur un cône consiste à effectuer une régression multiple ordinaire de **z** sur les variables indicatrices de l'ordre, définie par la matrice **Y** (n,n) suivante :

$$\begin{matrix} 1\;0\;0\;0\;0\;0\;0\;0\;0\;0\;0\;0 \\ 1\;1\;0\;0\;0\;0\;0\;0\;0\;0\;0\;0 \\ 1\;1\;1\;0\;0\;0\;0\;0\;0\;0\;0\;0 \\ 1\;1\;1\;1\;0\;0\;0\;0\;0\;0\;0\;0 \\ \ldots\ldots\ldots\ldots \\ \ldots\ldots\ldots\ldots \\ 1\;1\;1\;1\;1\;1\;1\;1\;1\;1\;1\;1 \end{matrix}$$

Si les coefficients sont tous positifs, sauf le premier, alors on a obtenu la projection sur le cône, sinon on élimine les colonnes correspondants à des coefficients négatifs ce qui revient à mettre ces coefficients à zéro, d'où des ex-aequo, et on recommence la régression multiple jusqu'à ce que tous les coefficients soient positifs ou nuls.

Un algorithme itératif, tel celui de SPAD-MC, consiste alors en une première étape de projection du vecteur non transformé des rangs sur W, donc des moindres carrés ordinaires, puis à projeter le vecteur des rangs estimés sur le cône, puis cette projection sur W etc.

L'analyse de variance avec transformation monotone de la réponse avait été proposée initialement par J.B.Kruskal pour éliminer des effets d'interaction dus à la non-linéarité de la relation entre réponse et facteurs : le cas classique est celui d'un modèle à effet multiplicatif Y= X1*X2 qui conduit à déclarer non nulle l'interaction entre X1 et X2 dans un modèle linéaire classique.

Une transformation logarithmique de Y rend le modèle additif sans interaction. Comme cette transformation est une transformation monotone particulière, on conçoit que la recherche de telles transformations aboutit dans de nombreux cas à éliminer les non-additivités éventuelles.

Quelle méthode d'estimation choisir, moindres carrés ordinaires ou régression monotone ? Il est clair que les transformations monotones sont mieux adaptés à la nature ordinale des données recueillies et au modèle additif postulé, mais peuvent conduire à des ajustements artificiels si le nombre de scénarios est insuffisant (risque d' »overfitting »).Il est donc recommandé d'examiner en premier lieu la solution des moindres carrés ordinaires avant de passer à la régression monotone : les résultats des MCO sont en général plus robustes.

3- INTERPRÉTATION DES RÉSULTATS

3.1 Cohérence

La reconstitution de l'ordre des préférences se mesure, soit avec le R^2, pour les moindres carrés ordinaires, soit avec le Tau de Kendall. Lorsque ces coefficients sont faibles ils traduisent soit une inadéquation du modèle d'utilités additives, soit une incohérence des réponses de l'enquêté. C'est en général cette dernière solution qui est retenue puisque l'on croit au modèle...

On éliminera alors les enquêtés pour lesquels le R^2 ou le Tau
est inférieur à un seuil fixé par l'utilisateur.

3.2 Hiérarchie des facteurs

Pour chaque enquêté on peut ordonner les facteurs ou variables explicatives. Le passage à une hiérarchie sur l'ensemble des enquêtés suppose l'homogénéité de ces derniers.

Puisqu'il s'agit en fait d'un modèle d'analyse de la variance à effets additifs on pense immédiatement à utiliser les sommes de carrés expliqués par chaque facteur. Il faut cependant être prudent car d'une part les sommes de carrés ont tendance à augmenter avec le nombre de modalités de chaque facteur, et d'autre part ces sommes ne sont additives que dans le cas particulier des plans orthogonaux.

Il faut alors utiliser des sommes de carrés de type II avec les tests F associés.

Une solution plus courante pour déterminer les importances relatives des différents facteurs consiste à utiliser les différences entre les utilités.

Plus précisément on cherchera tout d'abord les scénarios ou produits ayant des scores extrêmes parmi tous les scénarios possibles et non seulement parmi ceux proposés au jugement.

Ainsi dans l'exemple présenté plus haut le scénario préféré serait

A1 B3 C4 D3 qui donne le score le plus faible 0 - 4.6 - 3.3 - 0.3 = -8.2 et le scénario le moins accepté serait A3 B1 C1 D4 qui donne le score le plus fort 8 + 0 + 0 + 2.8 = 10.8

L'amplitude des scores est donc de 10.8 - (-8.2) = 19.

Le facteur A dont les utilités varient de 0 à 8 a une « importance » de 8/19 soit 42%.

Le facteur B dont les utilités varient de 0 à -4.6 une importance de 4.6/19 = 24%.

Le facteur C dont les utilités varient de 0 à -3.3 une importance de 3.3/19 = 17% et le facteur D dont les utilités varient de -0.3 à 2.8 une importance de 3.1/19 = 16%.

3.3 Les moyennes et leurs dangers

Il est assez fréquent de donner des utilités moyennes sur l'ensemble de l'échantillon et d'effectuer ensuite les calculs d'importance définis plus haut.

Cette pratique doit être effectuée avec précaution, d'autant que résumer par des moyennes va à l'encontre de l'esprit du trade-off qui consiste à suivre au plus près les préférences de chaque individu.

Il faut tout d'abord s'assurer de l'homogénéité de l'échantillon : un trade-off concernant des options non-fumeur dans les trains donnait évidemment des utilités opposées selon que l'enquêté était ou non fumeur. Une moyenne d'utilité sur l'ensemble n'aurait eu aucun sens et l'analyse devait être effectuée pour chaque groupe.

De même il est préférable d'effectuer des calculs d'importance moyenne en moyennant les importances des facteurs obtenues pour chaque enquêté plutôt que de les calculer sur les utilités moyennes.

3.4 Typologies et trade-off

Compte tenu de ce qui précède, il est indispensable d'analyser les résultats selon des groupes homogènes d'individus et donc de recourir à des typologies. Deux stratégies sont alors possibles :

- soit on utilise une typologie a priori établie sur des variables externes au trade-off et on vérifie la cohérence des choix à l'intérieur de chacun de ses groupes.
- soit on bâtit une typologie à partir du tableau donnant pour chaque enquêté les utilités des modalités des facteurs, ou encore les importances des facteurs qui sont souvent plus stables, typologie que l'on explique ensuite par des variables supplémentaires socio-démographiques par exemple.

Il convient donc d'associer étroitement les techniques de l'analyse des données aux résultats du trade-off.

4- SIMULATION

4.1 Nouveaux produits et parts de marché.

Un des principaux intérêts du trade-off réside dans la possibilité d'estimer des préférences vis à vis de produits non proposés dans l'enquête grâce au modèle d'utilités additives. Ainsi l'exemple sur les services en petit électro-ménager conduit à un nombre potentiel de produits de 3x3x4x4 = 144. Supposons que le choix final se réduise aux scénarios A1 B3 C2 D1 et A2 B1 C1 D4 ; on peut alors calculer pour chaque enquêté les scores de ces deux produits et en déduire la proportion de ceux qui préfèrent le premier au second, d'où une estimation des parts de marché de ces produits. Cette méthode est dite de l'utilité maximale.

Il existe deux autres méthodes classiques qui évitent la brutalité de ce choix lorsque les utilités sont proches :

le modèle logistique où la probabilité de choix d'un produit d'utilité U_i est proportionnelle à $\exp(U_i)$

le modèle de Bradley Terry Luce où cette probabilité est proportionnelle à U_i (ce qui ne peut s'appliquer que pour des utilités positives).

On estimera alors les parts de marché par la moyenne des probabilités sur l'ensemble des interviewés.

4.2 Acceptabilité et courbe d'intention d'achat.

Classer tous les scénarios ou produits proposés est une contrainte due au recueil des données ; rien n'indique que le produit classé en $8^{ème}$ position serait acheté par le consommateur ; il faut donc lui poser explicitement la question.

Si on dispose pour chaque produit de la variable oui-non, produit acceptable ou pas, on peut étudier le pourcentage d'acceptation ou de rejet de chaque produit. Si les conditions d'homogénéité sont respectées, on tracera le nuage de points mettant en abscisse le score moyen d'utilité *s* de chaque produit et en ordonnée le taux d'intention d'achat *p*; en général un ajustement par une fonction logistique

$$p = \exp(a + bs)/(1 + \exp(a + bs))$$

donne de bons résultats et permet de positionner les produits non testés.

5- CONSTRUCTION DES PRODUITS PAR PLANS D'EXPÉRIENCE

5.1 Introduction

Si un produit se définit par p facteurs à m_1 m_2... m_p catégories, il y a (m, x m_2 x ... x m_p) produits possibles. Les présenter tous constitue un plan factoriel complet. Le nombre total de produits rend en général irréaliste cette démarche. Ainsi l'exemple réel suivant qui consistait à étudier des services aériens caractérisés par les 5 facteurs :

Classe (affaire, économique) ;
Horaires (h1,h2) ;
Accès aux salons (oui,non) ;
Repas (chaud, froid) ;
Service à bord (réduit,normal)

conduit à 2^5=32 scénarios possibles.

Le nombre de coefficients (les utilités) à estimer étant égal à $\Sigma(m_i - 1) = \Sigma m_i - p$, on proposera un nombre de produits, ou scénarios, légérement supérieur à ce nombre.
Comment choisir ces scénarios ? Il s'agit ici d'estimer le mieux possible les coefficients d'un modèle linéaire. La théorie des plans d'expériences permet de répondre à cette question. On utilisera lorsque cela est possible des fractions orthogonales du plan complet qui permettent de bien séparer les influences des facteurs. Sinon on utilisera en général des plans D-optimaux.

Utiliser des plans fractionnaires, ou à nombre minimal d'essais, ne permet pas d'estimer certaines interactions et conduit à se limiter aux modèles à effets purement additifs, ce qui n'est pas gênant pour le trade-off car c'est précisément le modèle retenu. De plus on sait que les transformations monotones sur la réponse, c'est à dire les classements, ont pour effet d'annuler certaines interactions ou effets multiplicatifs.

5.2- Plans factoriels fractionnaires à 2 niveaux par facteurs

Reprenons l'exemple du service aérien.

Le plan complet à 32 scénarios est le suivant :

	A	B	C	D	E
1	1	1	1	1	1
2	2	1	1	1	1
3	1	2	1	1	1
4	2	2	1	1	1
5	1	1	2	1	1
6	2	1	2	1	1
7	1	2	2	1	1
8	2	2	2	1	1
9	1	1	1	2	1
10	2	1	1	2	1
11	1	2	1	2	1
12	2	2	1	2	1
13	1	1	2	2	1
14	2	1	2	2	1
15	1	2	2	2	1
16	2	2	2	2	1
17	1	1	1	1	2
18	2	1	1	1	2
19	1	2	1	1	2
20	2	2	1	1	2
21	1	1	2	1	2
22	2	1	2	1	2
23	1	2	2	1	2
24	2	2	2	1	2
25	1	1	1	2	2
26	2	1	1	2	2
27	1	2	1	2	2
28	2	2	1	2	2
29	1	1	2	2	2
30	2	1	2	2	2
31	1	2	2	2	2
32	2	2	2	2	2

Ce plan permet d'estimer les effets principaux mais aussi les interactions jusqu'à l'ordre 4. Ici chacune des 16 combinaisons de ABCD est présente deux fois.

Il est facile d'obtenir une moitié orthogonale de ce plan avec 16 essais. Ainsi on choisira ici une confusion ou générateur E=ABCD.

FRACTION ½

	A	B	C	D	E
1	1	1	1	1	2
2	2	1	1	1	1
3	1	2	1	1	1
4	2	2	1	1	2
5	1	1	2	1	1
6	2	1	2	1	2
7	1	2	2	1	2
8	2	2	2	1	1
9	1	1	1	2	1
10	2	1	1	2	2
11	1	2	1	2	2
12	2	2	1	2	1
13	1	1	2	2	2
14	2	1	2	2	1
15	1	2	2	2	1
16	2	2	2	2	2

Cette façon de créer le plan assure les meilleurs résultats en ce que les effets principaux ne sont confondus qu'avec des interactions d'ordre 4 et les interactions d'ordre 2 avec celles d'ordre 3.

On peut encore diviser par deux le nombre d'essais en utilisant les mêmes principes : on part du plan complet associé à A B C et on rajoute les colonnes D et E avec les générateurs D = AB et E = AC.

FRACTION ¼

	A	B	C	D	E
1	1	1	1	2	2
2	2	1	1	1	1
3	1	2	1	1	2
4	2	2	1	2	1
5	1	1	2	2	1
6	2	1	2	1	2
7	1	2	2	1	1
8	2	2	2	2	2

Ce dernier plan est le plan orthogonal au sens strict, minimal pour notre prob lème.

Le plan $L_8\ 2^7$ est couramment utilisé :

	A	B	C	D	E	F	G
1	1	1	1	2	2	2	1
2	2	1	1	1	1	2	2
3	1	2	1	1	2	1	2
4	2	2	1	2	1	1	1
5	1	1	2	2	1	1	2
6	2	1	2	1	2	1	1
7	1	2	2	1	1	2	1
8	2	2	2	2	2	2	2

Au delà de huit facteurs si on veut pas passer soit à 16 essais, on utilisera de préférence un plan de Plackett et Burman à 12 essais, plan très classique, construit à l'aide des matrices d'Hadamard.

5.3- Plans fractionnaires à k niveaux par facteur

Le plan factoriel complet a donc k^p essais.

Considérons le cas de lessives caractérisées par trois facteurs à trois catégories :

Marque (m1, m2, m3)
Conditionnement (classique, micro, liquide)
Prix (p1, p2, p3)

Il y a donc 27 combinaisons possibles : on peut les réduire à 9 en utilisant un plan en carré latin :

Conditionnement \ Marque	m1	m2	m3
classique	p1	p2	p3
micro	p2	p3	p1
liquide	p3	p1	p2

Ce qui donne en clair :

classique	m1	p1
classique	m2	p2
classique	m3	p3
micro	m1	p2
micro	m2	p3
micro	m3	p1
liquide	m1	p3
liquide	m2	p1
liquide	m3	p2

On constate que ce plan est équilibré et orthogonal au sens strict : tous les couples de niveaux sont présents une fois et une seule. Il y a ici 3 essais de plus que le nombre de paramètres.

Les carrés latins se construisent aisément pour k quelconque mais sont de peu d'utilité en trade-off car k^2 est en général trop grand si k dépasse 4.

Cette construction peut se généraliser à quatre facteurs, ce sont les carrés gréco-latins : ainsi, si on ajoute le facteur *Assouplissant* (α,β,γ) dans l'exemple des lessives, on utilisera le plan suivant :

Cond\ Marque\Assouplissant

	m1	m2	m3
classique	p1 α	p2 β	p3 γ
micro	p2 γ	p3 α	p1 β
liquide	p3 β	p1 γ	p2 α

Ce dernier plan ne comporte qu'un scénario de plus que le nombre de paramètres.

5.4- Plans orthogonaux pour facteurs à nombres de niveaux différents

On ne peut donner aucun principe général de construction, et des travaux de recherches se poursuivent activement dans ce domaine.

Voici cependant quelques astuces permettant de résoudre de nombreuses situations :

Reprenons l'exemple des lessives et supposons qu'il n'y a plus que deux niveaux de prix au lieu de trois. On a donc à construire un plan avec deux facteurs à trois niveaux et un facteur à deux niveaux. Dans ce cas la théorie dit que le plan minimal orthogonal au sens strict, comporte 18 scénarios qui est le plus petit commun multiple de 3x3 et 3x2. Or il suffit de créer une troisième catégorie fictive de prix pour avoir un plan à 9 scénarios, le carré latin étudié précédemment.

On part donc de ce dernier et on remplace après-coup p3 par p2 (*collapsing* de niveaux). Le plan obtenu est alors orthogonal au sens large car p1 est présent trois fois et p2 six fois.

classique	m1	p1
classique	m2	p2
classique	m3	p2
micro	m1	p2
micro	m2	p2
micro	m3	p1
liquide	m1	p2
liquide	m2	p1
liquide	m3	p2

Voici un exemple voisin :

il s'agissait d'un aliment pour bétail décrit par 4 facteurs à 3 2 2 et 4 niveaux respectivement. Le plus petit commun multiple de 3x2, 2x2, 4x2 et 4x3 est 24 ce qui est excessif. On a pu trouver un plan orthogonal à 12 scénarios en utilisant un plan trouvé dans un ouvrage pour le cas de 5 facteurs, l'un à 3 niveaux et les 4 autres à 2 niveaux chacun :

	A	B	C	D	E
1	1	1	1	1	1
2	1	1	1	2	2
3	1	2	2	1	1
4	1	2	2	2	2
5	2	1	1	1	2
6	2	1	2	2	1
7	2	2	1	2	1
8	2	2	2	1	2
9	3	1	2	1	1
10	3	1	2	2	2
11	3	2	1	1	2
12	3	2	1	2	1

Si on remarque qu'un facteur à 4 niveaux (a, b, c, d) peut s'écrire comme le croisement de deux facteurs à 2 niveaux (11, 12, 21, 22) le plan précédent peut convenir en regroupant les colonnes C et D pour lesquelles on trouve précisément 3 fois chacune de ces configurations associées à chacun des niveaux de A, d'où le plan :

	A	B	CD	E
1	1	1	a	1
2	1	1	b	2
3	1	2	c	1
4	1	2	d	2
5	2	1	a	2
6	2	1	d	1
7	2	2	b	1
8	2	2	c	2
9	3	1	c	1
10	3	1	d	2
11	3	2	a	2
12	3	2	b	1

Ce plan, peu classique, comporte 5 scénarios de plus que le nombre de paramètres et n'a pas été utilisé car 12 était encore trop grand.

5.5- Plans optimaux

Une des contraintes essentielles dans ce type d'études est d'avoir un faible nombre de scénarios (16 est un maximum pour des raisons de passation) et il n'est pas toujours possible de trouver des plans orthogonaux.

Trouver les meilleurs plans d'expériences à nombre de scénarios fixés suppose que l'on a défini un critère d'optimalité et que l'on dispose d'un algorithme efficace de recherche.

Que **y** soit le vecteur des rangs de classements, ou un transformé monotone de ce dernier, l'optimalité ne change pas car il s'agit d'un modèle linéaire. Si **X** représente la matrice des variables explicatives qui sont ici les indicatrices des niveaux des facteurs (avec une indicatrice de moins que de niveaux pour assurer l'inversibilité de **X'X**) les utilités sont les solutions **b** des équations normales :

$$\mathbf{b} = (\mathbf{X'X})^{-1}\mathbf{X'y}$$

Les utilités sont estimées sans biais mais avec une variabilité qui est définie par la matrice de variance des résidus et qui vaut :

$$V(b) = \sigma^2 (X'X)^{-1}$$

On utilisera le plus souventle critère du déterminant minimal pour $(\mathbf{X'X})^{-1}$ et donc maximal pour **X'X** ; c'est à dire la *D-optimalité*.

Les algorithmes consistent à rechercher parmi l'ensemble des solutions candidates , en général en partant du plan factoriel complet, le ou les sous-ensembles de scénarios de taille fixée maximisant le déterminant de **X'X**.

Ainsi dans l'exemple des aliments pour bétail, le plan D-optimal à 9 scénarios a été trouvé à l'aide de la PROC OPTEX de SAS-QC :

	A	B	CD	E
1	1	1	c	1
2	1	2	b	2
3	2	1	d	2
4	2	2	a	2
5	2	2	b	1
6	3	1	b	2
7	3	1	a	1
8	3	2	c	2
9	3	2	d	1

Lorsqu'il existe un plan orthogonal, celui-ci étant optimal, la procédure le trouve, sinon on obtient un compromis en général efficace. On peut également inclure d'office certains scénarios et optimiser sur les scénarios complémentaires.

Remarquons pour finir que pour des facteurs qualitatifs, la numérotation des modalités est arbitraire. On peut donc renuméroter à sa guise les catégories de chaque facteur, une fois le plan obtenu, ce qui permet d'écarter des scénarios irréalistes ou sans intérêt (un produit parfait qui serait classé en premier par tous) et d'inclure éventuellement des scénarios spécifiés par le client.

Références

Agha A.K. ,1991, *Régression et analyse canonique sous contraintes linéaires*, Thèse Université Paris-Dauphine

Evrard, Pras, Roux, 1993, « *Market* », Nathan

Green P.E., Srinivasan V. , 1990, *Conjoint analysis in marketing : new developments with implications for research and practice*, Journal of Marketing, 3-19

Kruskal J.B.,1965, *Analysis of factorial experiments by estimating monotone transformations of the data*, JRSS, B, 27, 251-263

SAS Institute Inc,1993, SAS ® Technical Report R-109, *Conjoint Analysis Examples*, Cary, NC.

Tenenhaus M., 1988, *Canonical analysis of two polyedral convex cones*, Psychometrika, 53, 4, 503-524