LES PLANS FRACTIONNAIRES.

A. Kobilinsky, Août 2001

Laboratoire de Biométrie.
INRA. Route de St-Cyr.
F78026 VERSAILLES Cedex.

1 Fractions de plans régulières pour facteurs à deux niveaux

1.1 Introduction

Au départ d'étude où l'on étudie l'influence de nombreux facteurs sur certaines variables, la première question est souvent de déterminer les facteurs réellement actifs. Pour éviter de multiplier les observations, on commence alors par étudier chaque facteur à 2 niveaux seulement. Mais même avec ce nombre minimum de 2 niveaux pour chaque facteur, le nombre de traitements définis par les combinaisons des niveaux des facteurs grimpe très vite. Ce nombre de traitement est 2^n s'il y a n facteurs, soit 512 pour $n = 9$ facteurs, 1024 pour 10 facteurs, etc ... Il devient donc très vite impossible d'expérimenter le *plan factoriel complet* formé par l'ensemble de ces traitements et on est conduit à sélectionner une fraction de taille compatible avec les possibilités expérimentales.

Il y a plusieurs possibilités pour former de telles fractions.

Il existe des *méthodes algorithmiques* qui cherchent automatiquement une fraction optimale pour un critère ad hoc. Ces méthodes sont apparemment très simples d'emploi. On précise le modèle reliant la réponse aux facteurs de variations étudiés, le nombre d'expériences réalisables et le programme fournit généralement un plan optimum. En fait la simplicité apparente cache des difficultés certaines. Ainsi le plan peut être très dépendant du modèle retenu. Comme le modèle est généralement ignoré, on le spécifie sous forme d'une approximation du type développement de Taylor. Pour être tranquille, on peut retenir une approximation de degré élevé, mais le nombre de données requises pour estimer tous les paramètres peut être alors hors d'atteinte. Si au contraire, on retient un modèle trop restreint, le critère utilisé peut être parfaitement inadéquat parce qu'il ne prend pas en compte l'incertitude sur le modèle.

On peut s'attendre à des progrès substantiels dans l'utilisation de ces méthodes algorithmiques. Mais il est vraisemblable qu'elles ne se substitueront jamais complétement aux méthodes algébriques conduisant aux plans dits réguliers.

Ces plans comprennent en particulier toutes les fractions de plans pour facteurs à 2 niveaux figurant dans les manuels les plus classiques sur les plans d'expérience [John P.W.M, 1971]. Ces plans ont des propriétés d'optimalité tout à fait remarquables [Kobilinsky, 1990]. Ils fournissent des plans utilisables même quand le modèle retenu a plus de paramètres que de données (plans de résolution 4). C'est donc à la maîtrise de ces plans que seront consacrés les paragraphes qui suivent. Nous considérerons d'abord le cas des facteurs à 2 niveaux. Nous avons vu son importance pratique. Sur le plan mathématique, il s'avère qu'il est certainement le cas le mieux connu et le plus facile à présenter.

1.2 Modèle dans le cas de deux facteurs

Pour simplifier la présentation, on s'appuiera sur l'exemple d'une étude visant à préciser l'influence des deux facteurs température (A) et pH (B) sur le taux de croissance y d'une bactérie.

Dans un premier temps, on peut décider d'étudier seulement les deux températures 20°, 30° et les deux pH 7.5, 8.5. Les traitements sont alors les 4 couples $\{(20°,7.5), (20°,8.5), (30°,7.5), (30°,8.5)\}$. Pour expliciter les constructions de plan, il est commode de coder -1 et 1 les deux niveaux de chaque facteur. Le pH et la température étant des facteurs qunatitatifs, on peut choisir de coder -1 le niveau bas, c'est à dire 20° pour la température, 7.5 pour le pH. C'est ce que nous ferons ici. Mais il faut être conscient que ce choix n'a rien d'impératif et que l'on peut tout aussi bien, dans ce qui suit, tirer au hasard quand il y a deux niveaux celui qui est codé 1 et celui qui est codé -1.

Avec ce codage, si on utilise pour désigner les niveaux codés d'un facteur la même lettre que celle désignant ce facteur, les traitements sont les couples (A,B) où A et B prennent les valeurs -1, 1, soit : $\{(-1,-1), (-1,1), (1,-1), (1,1)\}$.

Lorsqu'on expérimente avec un de ces traitements (A,B), la réponse y fluctue autour d'une valeur moyenne notée ici $\tau(A,B)$. Cette valeur est appelée *effet du traitement* (A,B), ou encore *réponse théorique* du traitement (A,B). Les statisticiens l'appelle *espérance* de la réponse y pour le traitement (A,B) et la note $E(y|A,B)$:

$$E(y|A,B) = \tau(A,B) \ .$$

La variance des réponses obtenue avec un traitement fixe (A,B) est aussi un élément important pour caractériser la façon dont y dépend des deux facteurs. Une hypothèse souvent faite est que cette variance ne dépend pas du traitement (A,B) considéré, qu'elle est donc égale à une constante σ^2 quel que soit le traitement considéré :

$$\mathrm{var}(y|A,B) = \sigma^2 \ . \tag{1}$$

Il y a deux bonnes raisons pour faire une telle hypothèse. D'une part, elle est très plausible si les niveaux des facteurs considérés ne sont pas trop éloignés, c'est à dire dans cet exemple si on s'attend à ce que le comportement du microorganisme soit d'une même nature aux deux températures et pH considérés. D'autre part, la non validité de cette hypothèse affecte généralement peu les comparaisons entre moyennes effectuées. Et compte tenu du nombre prohibitif de répétitions des traitements nécessaire pour évaluer correctement les variances pour chaque traitement, cette dernière raison explique que cette hypothèse dite *d'homoscédasticité* soit souvent adoptée.

Lorsqu'on expérimente avec le traitement (A,B), la différence $\varepsilon = y - \tau(A,B)$ est le *bruit* ou *erreur* dû à la fluctuation de la réponse. A cause de cette erreur, les réponses théoriques $\tau(A,B)$ ne peuvent jamais être obtenues par l'expérience et on ne peut en obtenir que des estimations que l'on note souvent en ajoutant un chapeau sur le τ. Ainsi $\hat{\tau}(A,B)$ désigne l'estimation de l'effet traitement $\tau(A,B)$. L'estimation la plus naturelle est la moyenne des réponses obtenues avec le traitement (A,B). Si ces réponses sont notées $y(A,B,1)$, $y(A,B,2)$, ... , $y(A,B,r)$ où r est le nombre de répétitions du traitement (A,B), l'estimation naturelle est donc :

$$\hat{\tau}(A,B) = y(A,B,\bullet) = \frac{y(A,B,1) + \cdots + y(A,B,r)}{r} \ .$$

Le point à la place du 3$^{\text{ème}}$ indice est une notation couramment utilisée pour indiquer que l'on fait la moyenne sur cet indice.

Cette estimation naturelle ne peut cependant pas être utilisée dans les fractions de plan puisque tous les traitements ne sont pas expérimentés. Nous verrons ultérieurement quelles hypothèses et procédures sont utilisées dans ces fractions pour estimer les réponses théoriques.

Sous l'hypothèse d'homoscédasticité, la variance σ^2 de la réponse obtenue avec un traitement fixé est aussi appelée *variance d'erreur* ou *variance résiduelle*. Cette dernière dénomination vient de ce que cette variance est celle qui reste quand on soustrait l'effet traitement à chaque réponse.

1.3 Effets factoriels

Dans les études portant sur plusieurs facteurs, on ne cherche pas seulement à préciser les effets de chaque traitement, mais aussi et surtout à dire quels sont les facteurs importants, quel est leur mode d'action et comment ils interagissent. Pour cela, on évalue des quantités synthétiques, appelées *effets factoriels* qui se déduisent des effets des traitements.

Le cas de l'exemple avec les deux facteurs suffit pour expliciter le mode de formation de ces quantités.

A partir des 4 réponses théoriques on définit la moyenne générale, les effets principaux des facteurs A et B et leur interaction par les formules suivantes

$$\begin{array}{llllll}
\text{moyenne générale} & : & e(\mathbf{1}) & = \mu & = & \dfrac{\tau(-1,-1)+\tau(-1,1)+\tau(1,-1)+\tau(1,1)}{4} \\[2ex]
\text{effet principal } A & : & e(A) & = \alpha & = & \dfrac{-\tau(-1,-1)-\tau(-1,1)+\tau(1,-1)+\tau(1,1)}{4} \\[2ex]
\text{effet principal } B & : & e(B) & = \beta & = & \dfrac{-\tau(-1,-1)+\tau(-1,1)-\tau(1,-1)+\tau(1,1)}{4} \\[2ex]
\text{interaction } A.B & : & e(AB) & = \gamma & = & \dfrac{\tau(-1,-1)-\tau(-1,1)-\tau(1,-1)+\tau(1,1)}{4}
\end{array} \qquad (2)$$

La *moyenne générale*, désignée par $e(\mathbf{1})$ ou par la lettre grecque μ, est la moyenne des 4 réponses théoriques.

L'*effet principal* du facteur A, noté par $e(A)$ ou par α, est la demi-différence entre la moyenne des réponses théoriques au niveau codé 1 de A et la moyenne au niveau codé -1:

$$e(A) = \frac{1}{2}\left[\frac{\tau(1,-1)+\tau(1,1)}{2} - \frac{\tau(-1,-1)+\tau(-1,1)}{2}\right] .$$

De même l'effet principal de B, boté $e(B)$ ou β, est la demi-différence entre les moyennes aux deux niveaux de B.

Enfin l'*interaction* $A.B$, notée $e(AB)$ ou γ, s'obtient à un coefficient près, comme la demi-différence entre les effets du facteur B aux deux niveaux de A. L'effet de B au niveau 1 de A est par définition la demi-différence $e(B|A=1) = [\tau(1,1) - \tau(1,-1)]/2$. De même l'effet de B au niveau -1 de A est $e(B|A=-1) = [\tau(-1,1) - \tau(-1,-1)]/2$. Et on voit que la demi-différence entre ces deux effets

$$\frac{e(B|A=1)-e(B|A=-1)}{2}=\frac{\tau(1,1)-\tau(1,-1)-\tau(-1,1)+\tau(-1,-1)}{4}$$

est bien la quantité $e(AB)$ figurant dans (2).

Il n'y a pas d'interaction quand $e(B|A=1)=e(B|A=-1)$, c'est à dire quand l'effet de B ne dépend pas du niveau de A. Et lorsque c'est le cas, on vérifie immédiatement que l'effet commun est l'effet principal de B : $e(B)=e(B|A=1)=e(B|A=-1)$.

La définition que l'on vient de donner introduit une dissymétrie entre A et B. En fait, la définition est symétrique en A et B et on peut également définir l'interaction comme la demi-différence entre les effets du facteur A aux deux niveaux de B :

$$e(AB)=\frac{1}{2}\left[e(A|B=1)-e(A|B=-1)\right]=\frac{1}{2}\left[\frac{\tau(1,1)-\tau(-1,1)}{2}-\frac{\tau(1,-1)-\tau(-1,-1)}{2}\right].$$

L'interaction est nulle si l'effet de A ne dépend pas du niveau de B, et dans ce cas l'effet commun est l'effet principal de A : $e(A)=e(A|B=1)=e(A|B=-1)$.

L'effet principal du facteur A peut aussi être défini comme l'effet moyen de A, la moyenne étant effectuée sur les niveaux de B:

$$e(A)=\frac{e(A|B=1)+e(A|B=-1)}{2}=\frac{1}{2}\left[\frac{\tau(1,1)-\tau(-1,1)}{2}+\frac{\tau(1,-1)-\tau(-1,-1)}{2}\right]$$

Cela n'a pas d'intérêt de faire une telle moyenne si les deux effets dont on fait la moyenne sont très différents, c'est à dire si il y a une très forte interaction. Mais dans bien des cas, l'interaction est relativement faible. Les effets de A aux deux niveaux de B, sans être rigoureusement identiques, sont suffisemment voisins pour que cela ait un sens d'en faire la moyenne.

Les 4 effets définis par (2) sont appelés *effets factoriels* des facteurs A, B. Il est assez usuel de les noter avec des lettres grecques. Mais la notation $e()$ avec entre parenthèse le produit des facteurs associés (i.e. ($e(\mathbf{1})$, $e(A)$, $e(B)$, $e(AB)$) a l'avantage d'être mnémotechnique et facile à généraliser.

On notera que la définition (2) peut être synthétisée par la formule

$$e(A^aB^b)=\frac{1}{4}\sum_{A,B}A^aB^b\tau(A,B)\ . \tag{3}$$

Par exemple si (a,b) prend la valeur $(1,0)$, $A^aB^b=A$ et (3) donne

$$e(A)=\frac{1}{4}\sum_{A,B}A\tau(A,B)=\frac{1}{4}[-\tau(-1,-1)-\tau(-1,1)+\tau(1,-1)+\tau(1,1)]\ .$$

Pour concrétiser d'avantage, imaginons que les réponses théoriques $\tau(A,B)$ sont connues et données par:

$$\tau(-1,-1)=2,\quad \tau(-1,1)=4,\quad \tau(1,-1)=14,\quad \tau(1,1)=8\ . \tag{4}$$

Cette situation est illustrée par la figure 1. Les effets factoriels sont alors les suivants

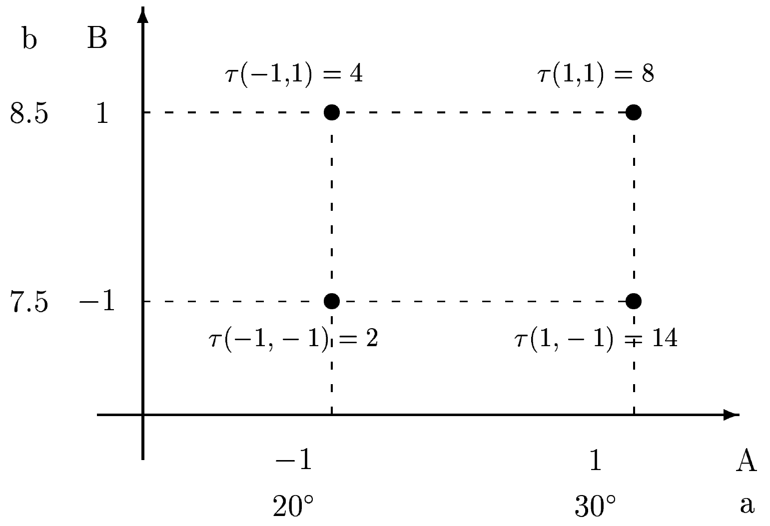

FIG. 1 – *Points et réponses théoriques d'un plan factoriel complet* 2^2

Les minuscules a et b dénotent les niveaux réels, A et B les niveaux codés. La figure donne les valeurs des réponses théoriques $\tau(A,B)$ imaginées pour concrétiser. En pratique ces valeurs ne peuvent qu'être approchées à partir des réponses observées

$$\begin{array}{lcccc}
\text{moyenne générale} & \mu & = & 7 & = \dfrac{2+4+14+8}{4} \\
\text{effet principal } A & \alpha & = & 4 & = \dfrac{11-3}{2} \\
\text{effet principal } B & \beta & = & -1 & = \dfrac{6-8}{2} \\
\text{interaction } A.B & \gamma & = & -2 & = \dfrac{2-6}{2}
\end{array} \tag{5}$$

Les figures 2 et 3 en donnent une représentation graphique dans laquelle la réponse théorique τ figure en ordonnée, le niveau de A figure en abscisse et le niveau de B est indiqué par un symbole ($\bullet$ ou $\circ$). Le tableau 1 fait le lien entre les égalités de définition générale (2) et ce cas particulier.

En pratique les réponses théoriques ne sont pas connues. On ne peut que les approcher en répétant un grand nombre de fois l'expérimentation de chaque traitement. Les effets factoriels qui s'en déduisent ne sont donc pas connus non plus. Il est commode cependant d'imaginer des valeurs pour se faire des représentations plus concrètes. Mais il faut bien distinguer ces quantités théoriques, inaccessibles par l'expérimentation, des estimations que l'on en fera ultérieurement à partir des expériences.

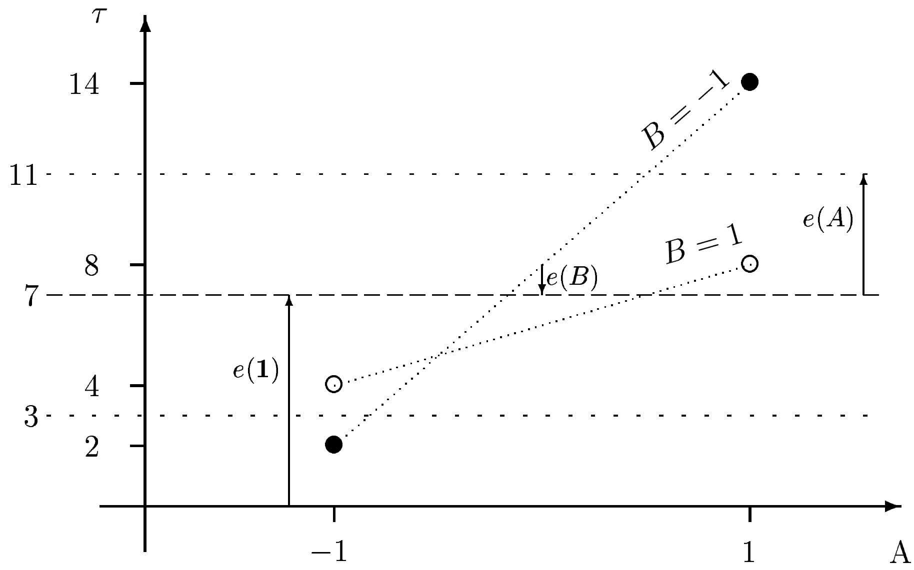

FIG. 2 – *Représentation graphique des effets principaux*

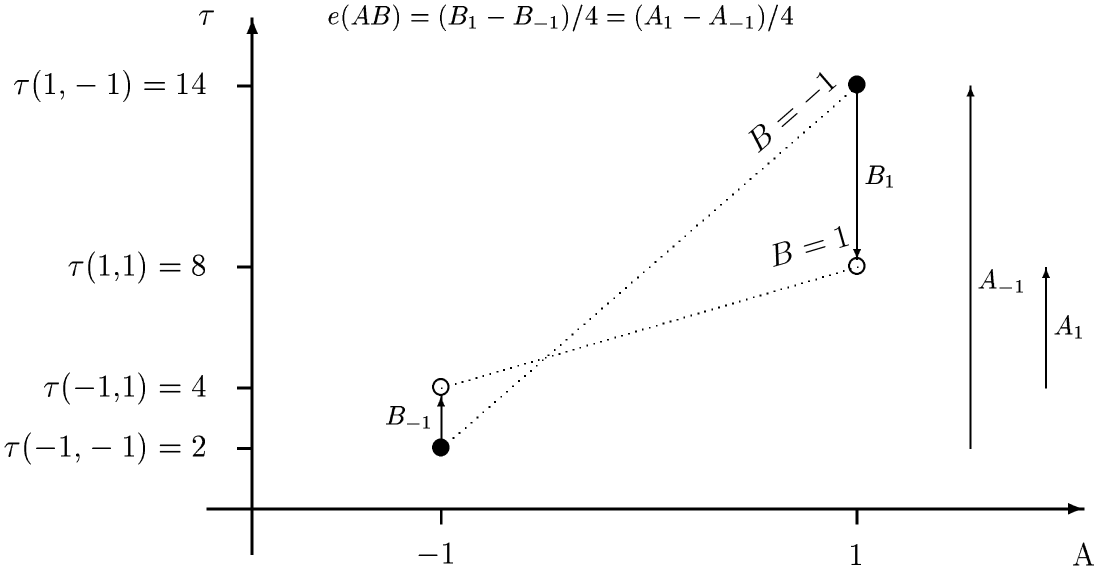

FIG. 3 – *Eléments pour le calcul de l'interaction*

On a noté ici A_1 la différence entre les réponses théoriques aux deux niveaux de A quand B vaut 1: $A_1 = \tau(1{,}1) - \tau(-1{,}1)$. Les notations B_1, A_{-1}, B_{-1} sont similaires.

$$
\begin{array}{|llcll|}
\hline
\text{moyenne générale :} & e(\mathbf{1}) & = & \dfrac{\tau(-1,-1)+\tau(-1,1)+\tau(1,-1)+\tau(1,1)}{4} & \\
 & & = & \dfrac{2+4+14+8}{4} & = 7 \\
\text{effet principal } A : & e(A) & = & \dfrac{1}{2}\left[\dfrac{\tau(1,-1)+\tau(1,1)}{2}-\dfrac{\tau(-1,-1)+\tau(-1,1)}{2}\right] & \\
 & & = & \dfrac{1}{2}\left[\dfrac{14+8}{2}-\dfrac{2+4}{2}\right] & = 4 \\
\text{effet principal } B : & e(B) & = & \dfrac{1}{2}\left[\dfrac{\tau(-1,1)+\tau(1,1)}{2}-\dfrac{\tau(-1,-1)+\tau(1,-1)}{2}\right] & \\
 & & = & \dfrac{1}{2}\left[\dfrac{4+8}{2}-\dfrac{2+14}{2}\right] & = -1 \\
\text{interaction } A.B : & e(AB) & = & \dfrac{1}{2}\left[\dfrac{\tau(1,1)-\tau(-1,1)}{2}-\dfrac{\tau(1,-1)-\tau(-1,-1)}{2}\right] & \\
 & & = & \dfrac{1}{2}\left[\dfrac{8-4}{2}-\dfrac{14-2}{2}\right] & = -2 \\
\hline
\end{array}
$$

TAB. 1 – *Calcul des effets factoriels à partir des effets traitements donné par (4)*

1.4 Modèle sur les espérances

Les espérances de y pour les 4 traitements (les effets traitement) s'exprime en fonction des effets factoriels par la relation

$$\tau(A, B) = e(\mathbf{1}) + A\ e(A) + B\ e(B) + AB\ e(AB)$$

qui donne les égalités

$$
\begin{array}{rcl}
\tau(-1,-1) &=& e(\mathbf{1}) - e(A) - e(B) + e(AB), \\
\tau(-1,\ 1) &=& e(\mathbf{1}) - e(A) + e(B) - e(AB), \\
\tau(\ 1,-1) &=& e(\mathbf{1}) + e(A) - e(B) - e(AB), \\
\tau(\ 1,\ 1) &=& e(\mathbf{1}) + e(A) + e(B) + e(AB).
\end{array}
\tag{6}
$$

Ces égalités sont faciles à vérifier en partant directement des définitions (2) : si on multiplie chaque ligne par le coefficient ad hoc et que l'on effectue la somme, on retrouve l'effet traitement concerné.

La relation entre les effets traitements et les effets factoriels s'exprime aussi commodément sous forme matricielle. Pour cela on forme le tableau 2 dont les colonnes A, B contiennent les niveaux codés des facteurs A et B, la colonne AB est le produit des colonnes A, B et la colonne $\mathbf{1}$ a toutes ses coordonnées égales à 1.

On vérifie immédiatement la règle suivante. Les effets factoriels définis en (2) s'obtiennent en multipliant chaque effet traitement à gauche par le coefficient de la colonne associée à cet effet, en sommant ces produits puis en divisant par 4.

$$
\begin{array}{c}
e' = \quad [e(\mathbf{1}) \quad e(A) \quad e(B) \quad e(AB)] \\[1em]
\begin{array}{cc}
\underbrace{\begin{array}{l}
\tau(\;\;A\;\;B) \\[1em]
\tau(-1,-1) \\
\tau(-1,\;\;1) \\
\tau(\;\;1,-1) \\
\tau(\;\;1,\;\;1)
\end{array}}_{\tau}
&
\underbrace{\begin{array}{rrrr}
\mathbf{1} & A & B & AB \\[1em]
1 & -1 & -1 & 1 \\
1 & -1 & 1 & -1 \\
1 & 1 & -1 & -1 \\
1 & 1 & 1 & 1
\end{array}}_{Z}
\end{array}
\end{array}
$$

TAB. 2 – *Relations entre effets factoriels et effets des traitements*

Lecture :

- pour obtenir l'effet factoriel en haut d'une colonne, multiplier les coefficients de la colonne par les réponses théoriques à gauche, sommer et diviser par 4.
- pour obtenir l'effet traitement à gauche d'une ligne, sommer les produits des coefficients de la ligne par les effets factoriels en haut.

Cela conduit à l'égalité matricielle

$$
\underbrace{\begin{bmatrix} e(\mathbf{1}) \\ e(A) \\ e(B) \\ e(AB) \end{bmatrix}}_{e}
=
\underbrace{\frac{1}{4}\begin{bmatrix} 1 & 1 & 1 & 1 \\ -1 & -1 & 1 & 1 \\ -1 & 1 & -1 & 1 \\ 1 & -1 & -1 & 1 \end{bmatrix}}_{\frac{1}{4}Z'}
\underbrace{\begin{bmatrix} \tau(-1,-1) \\ \tau(-1,\;\;1) \\ \tau(\;\;1,-1) \\ \tau(\;\;1,\;\;1) \end{bmatrix}}_{\tau}
\tag{7}
$$

Dans cette égalité, e, τ désignent respectivement les vecteurs des effets factoriels et des effets traitements. Et Z est la matrice 4×4 donnée au tableau 2 qui a $\mathbf{1}$, A, B, AB pour colonnes.

Il est facile de vérifier que

$$
Z'Z = ZZ' = 4\,\mathbf{I}\ . \tag{8}
$$

(le tableau 3 donne les matrices Z', Z et leur produit sous une forme facilitant le calcul).

On a donc en prémultipliant (7) par Z

$$
e = \frac{1}{4}Z'\tau \Longrightarrow Ze = \frac{1}{4}ZZ'\tau = \tau\ .
$$

$$
\begin{array}{cc}
 & \left.\begin{bmatrix} 1 & -1 & -1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & 1 & 1 & 1 \end{bmatrix}\right\} Z \\
\underbrace{\begin{bmatrix} 1 & 1 & 1 & 1 \\ -1 & -1 & 1 & 1 \\ -1 & 1 & -1 & 1 \\ 1 & -1 & -1 & 1 \end{bmatrix}}_{Z'} & \begin{bmatrix} 4 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 4 \end{bmatrix}
\end{array}
$$

TAB. 3 – *Matrices Z', Z et produit $Z'Z$*

Cette dernière égalité s'écrit sous la forme développée suivante

$$
\underbrace{\begin{bmatrix} \tau(-1,-1) \\ \tau(-1,\ 1) \\ \tau(\ 1,-1) \\ \tau(\ 1,\ 1) \end{bmatrix}}_{\tau} = \underbrace{\begin{bmatrix} 1 & -1 & -1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & 1 & 1 & 1 \end{bmatrix}}_{Z} \underbrace{\begin{bmatrix} e(\mathbf{1}) \\ e(A) \\ e(B) \\ e(AB) \end{bmatrix}}_{e} \tag{9}
$$

$$\tau = Z \quad e$$

Elle permet d'exprimer les effets traitements en fonction des effets factoriels. Elle montre que les effets traitements à gauche du tableau 2 s'obtiennent en multipliant les coefficients de la ligne associée par les effets factoriels en haut des colonnes correspondantes et en sommant ces produits. Cette égalité (9) redonne aussi les expressions (6) introduites précédemment.

Comparabilité des effets factoriels.
De par la définition (2), la moyenne des 4 effets traitements est la moyenne générale $e(\mathbf{1})$. La dispersion de ces effets peut se mesurer par leur variance

$$
V = \frac{1}{4}\sum_{A,B}\big(\tau(A,B) - e(\mathbf{1})\big)^2 = \frac{1}{4}\Big[\big(\tau(-1,-1) - e(\mathbf{1})\big)^2 + \big(\tau(-1,1) - e(\mathbf{1})\big)^2 + \big(\tau(1,-1) - e(\mathbf{1})\big)^2 + \big(\tau(1,1) - e(\mathbf{1})\big)^2\Big] . \tag{10}
$$

Il n'est pas difficile en utilisant (8) et (9) de vérifier que cette variance se met sous la forme:

$$
V = e(A)^2 + e(B)^2 + e(AB)^2 . \tag{11}
$$

La contribution à cette variance de chaque effet est donc égale à son carré, ce qui montre qu'avec la définition (2), les effets factoriels sont directement comparables. Il n'en serait pas de même si on avait utilisé des coefficients différents de $1/4$. Par exemple, si on avait omis la division par 4 pour la ligne $e(AB)$, l'effet défini aurait été 4 fois plus grand et on aurait dû rajouter un coefficient $1/16$ devant $e(AB)^2$ dans (11). Ceci aurait rendu impossible la comparaison directe entre $e(AB)$ et les deux autres effets $e(A)$, $e(B)$.

1.5 Etude expérimentale

1.5.1 Le modèle linéaire

Dans l'exemple avec les deux facteurs température (A) et pH (B), supposons qu'on expérimente une fois chacun des 4 traitements. On dit alors qu'on réalise le plan factoriel complet sans répétition.

On peut désigner par $y(A,B)$ la réponse (taux de croissance) obtenue avec le traitement (A,B). Cette réponse obéit au modèle :

$$y(A,B) = \tau(A,B) + \varepsilon(A,B) \tag{12}$$

où l'erreur $\varepsilon(A,B)$ est d'espérance nulle, de variance σ^2 sous l'hypothèse d'homoscédasticité :

$$\mathrm{var}\big(\varepsilon(A,B)\big) = \sigma^2 \,.$$

Cette hypothèse signifie que si l'on réalisait en fait une multitude de répétitions de ce même traitement (A,B), la réponse oscillerait autour de la moyenne théorique $\tau(A,B)$ avec la variance σ^2. En remplaçant $\tau(A,B)$ par son expression en fonction des effets factoriels (6), on obtient:

$$y(\ A,\ B) = e(\mathbf{1}) + A\ e(A) + B\ e(B) + AB\ e(AB) + \varepsilon(\ A,\ B)$$

qui conduit aux 4 égalités

$$\begin{array}{lclclclcll}
y(-1,-1) & = & e(\mathbf{1}) & - & e(A) & - & e(B) & + & e(AB) & +\varepsilon(-1,-1), \\
y(-1,\ \ 1) & = & e(\mathbf{1}) & - & e(A) & + & e(B) & - & e(AB) & +\varepsilon(-1,\ \ 1), \\
y(\ \ 1,-1) & = & e(\mathbf{1}) & + & e(A) & - & e(B) & - & e(AB) & +\varepsilon(\ \ 1,-1), \\
y(\ \ 1,\ \ 1) & = & e(\mathbf{1}) & + & e(A) & + & e(B) & + & e(AB) & +\varepsilon(\ \ 1,\ \ 1).
\end{array} \tag{13}$$

Ce système de 4 égalités est ce qu'on appelle le *modèle linéaire* décrivant les 4 réponses du plan. Il se réécrit sous la forme matricielle

$$\underbrace{\begin{bmatrix} y(-1,-1) \\ y(-1,\ \ 1) \\ y(\ \ 1,-1) \\ y(\ \ 1,\ \ 1) \end{bmatrix}}_{y} = \underbrace{\begin{bmatrix} 1 & -1 & -1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & 1 & 1 & 1 \end{bmatrix}}_{Z} \underbrace{\begin{bmatrix} e(\mathbf{1}) \\ e(A) \\ e(B) \\ e(AB) \end{bmatrix}}_{e} + \underbrace{\begin{bmatrix} \varepsilon(-1,-1) \\ \varepsilon(-1,\ \ 1) \\ \varepsilon(\ \ 1,-1) \\ \varepsilon(\ \ 1,\ \ 1) \end{bmatrix}}_{\varepsilon} \tag{14}$$

où e, Z ont la même signification qu'auparavant, y et ε sont le vecteur des observations et le vecteur des erreurs. Comme auparavant, les colonnes 2 et 3 de Z, associées aux effets $e(A)$, $e(B)$, donnent respectivement les niveaux de A et B pour les observations correspondantes, la colonne 1 associée à $e(\mathbf{1})$ a ses coordonnées toutes égales à 1 et la colonne 4 associée à AB est le produit des colonnes associées à A et B.

Cette dernière forme de modèle est bien connue. L'estimateur des moindres carrés du vecteur e des effets factoriels est donné par

$$\hat{e} = (Z'Z)^{-1}Z'y$$

soit en tenant compte de l'égalité (8)

$$\hat{e} = \frac{1}{4} Z' y \ .$$

On voit que, pour estimer le vecteur e des effets factoriels, il suffit dans ce cas de remplacer, dans l'expression (7) qui le définit ($e = Z'\tau/4$), les réponses théoriques $\tau(A,B)$ par les réponses observées $y(A,B)$ correspondantes. Ceci est vrai parce qu'il s'agit d'un plan complet sans répétition. On verra qu'on a un résultat semblable avec les fractions de plan régulières introduites plus loin. Mais si les traitements ne sont pas equirépétés, ce résultat devient faux.

Sous l'hypothèse d'homoscédasticité (1), on sait que la matrice de variance de l'estimateur $\hat{e}$ est

$$\mathrm{var}(\hat{e}) = \frac{\sigma^2}{4} \mathbf{I}_4 \ .$$

Rappelons que $\mathbf{I}_4$ est la matrice identité 4×4:

$$\mathbf{I}_4 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \ .$$

Autrement dit, avec ce plan factoriel complet, les effets factoriels estimés sont non corrélés et ont la même variance $\sigma^2/4$ qui est la variance résiduelle divisée par le nombre d'observations 4. Ce dernier résultat est évidemment lié à la façon dont sont définis les effets factoriels. Ainsi, si la division par 4 avait été omise dans (2), ligne $e(AB)$, on aurait trouvé un écart-type 4 fois plus grand. Avec ce qui a été dit à la fin du paragraphe 1.4, ceci milite pour choisir les coefficients dans la définition des effets factoriels comme cela est fait dans (2).

1.5.2 Intervalle de confiance des effets factoriels

On a dans cet exemple 4 données et 4 paramètres $e(\mathbf{1})$, $e(A)$, $e(B)$, $e(AB)$ à estimer. Il n'y a donc pas de degrés de liberté pour estimer la variance d'erreur et on ne peut donner des intervalles de confiance des paramètres (et tester leur nullité) que si cette variance d'erreur est déjà connue par des expériences effectuées auparavant dans des conditions similaires.

Supposons que ceci est le cas et que σ^2 est connu. Si la réponse y est supposée de distribution normale, chaque effet factoriel a une estimation de distribution normale d'où l'on déduit immédiatement un intervalle de confiance au seuil désiré. Par exemple, on a vu que l'effet principal de A est estimé par

$$\hat{e}(A) = \frac{-y(-1,-1) - y(-1,1) + y(1,-1) + y(1,1)}{4}$$

quantité où les réponses observées $y(A,B)$ remplace les réponses théoriques $\tau(A,B)$ figurant dans la définition (2) de $e(A)$. Cet estimateur $\hat{e}(A)$ a $e(A)$ pour espérance et $\sigma^2/4$ comme variance, donc $\sigma/2$ pour écart-type. Par suite $[\hat{e}(A) - e(A)]/2\sigma$ suit une loi normale d'espérance nulle, écart-type 1 dont on sait qu'elle est comprise entre -1.96 et 1.96 avec une probabilité 95%. On a donc avec une probabilité 95%

$$-1.96 \leq \frac{\hat{e}(A) - e(A)}{\sigma/2} \leq 1.96$$

et ces inégalités donne pour $e(A)$ l'intervalle de confiance à 95%

$$\hat{e}(A) - 1.96\sigma/2 \leq e(A) \leq \hat{e}(A) + 1.96\sigma/2 \ .$$

L'exemple qui précède est un cas d'école. Si une étude de cette sorte devait être réalisée, on répéterait chaque expérience au moins 2 fois. Avec deux répétitions de chacun des 4 traitements, il y aurait 8 réponses observées. Une fois estimé les 4 effets factoriels, il resterait donc $4 = 8 - 4$ degrés de liberté pour estimer la variance résiduelle σ^2. Le ratio précédent $[\hat{e}(A) - e(A)]/2\sigma$ où σ est inconnu serait remplacé par le ratio $[\hat{e}(A) - e(A)]/2\hat{\sigma}$ dont on sait qu'il suit la distribution du t de student centré avec 4 degrés de liberté. La table du t de student, où un programme ad hoc, montre qu'une telle statistique est comprise entre -2.78 et 2.78 avec une probabilité 95%. On aurait donc avec une probabilité 95%

$$-2.78 \leq [\hat{e}(A) - e(A)]/2\hat{\sigma} \leq 2.78$$

et l'intervalle de confiance à 95% de $e(A)$ serait

$$\hat{e}(A) - 2.78\hat{\sigma}/2 \leq e(A) \leq \hat{e}(A) + 2.78\hat{\sigma}/2 \ .$$

1.5.3 Exemple d'analyse

Le plan factoriel complet sans répétition pour l'exemple avec la température et le pH figure dans le fichier EX2P2.DAT, celui avec 2 répétitions dans le fichier EX2P2R.DAT. Pour que les choses soient claires, les deux facteurs ont été noté T et pH, non A et B, et ce sont leurs niveaux réels 20°, 30° pour T, 7.5, 8.5 pour le pH qui ont été reportées dans le fichier.

La réponse y a été simulée dans chaque cas en rajoutant aux réponses théoriques donnée en (4) une erreur simulée par tirage aléatoire d'une loi normale centrée d'écart-type 2. A titre indicatif, on a aussi mis dans le fichier une réponse exactement égale à la réponse théorique.

Les deux fichiers EX2P2.DAT, EX2P2R.DAT sont sous une forme directement analysable par ANALYS. Ce sont des fichiers texte comportant sur la première ligne les libellés des deux facteurs T, pH puis ceux des deux variables *tau*, y avec un # séparant les facteurs des variables. Les lignes suivantes donnent les valeurs de ces facteurs et variables dans le même ordre. Il y en a autant que d'unités expérimentales.

On a fait une analyse avec T et pH pris comme facteurs qualitatifs, donc en faisant abstraction du caractère quantitatif de ces deux facteurs. Le modèle $T.\,pH$ choisi conduit automatiquement dans ANALYS à une reparamétrisation par les effets factoriels $e(\mathbf{1})$, $e(T)$, $e(pH)$ et $e(T.\,pH)$.

L'analyse du plan complet sans répétition EX2P2.DAT donne exclusivement les estimations de ces effets. La correspondance entre les niveaux du fichier et les niveaux codés

$-1, 1$ utilisés pour définir les effets est indiquée en tête des sorties (sous le titre *Contrastes pour facteur ...*). L'analyse de la variable *tau* égale à la réponse théorique redonne les effets factoriels.

L'analyse du plan avec deux répétitions permet d'obtenir l'estimation $\hat{\sigma}^2$ de la variance résiduelle et permet donc de donner des intervalles de confiance des effets factoriels et de les tester. Les données et une partie des résultats de l'analyse effectuée sont reproduit dans le tableau 4. On trouve que l'effet principal $e(T)$ et l'interaction $e(T.\,pH)$ sont significativement différent de 0 au seuil respectivement 1%, 5%. Les estimations correspondantes sont $\hat{e}(T) = 4.402$ et $\hat{e}(T.\,pH) = -3.198$.

Pour interpréter directement ces effets, le mieux est de se référer au modèle (6). Il montre que $e(T)$ et $e(T.\,pH)$ sont à rajouter ou retrancher selon que le produit entre parenthèse – T ou $T.\,pH$ – vaut 1 ou -1. Ainsi, le taux de croissance y augmente de 4.402 si $T = 1$, diminue de 4.402 si $T = -1$. Compte tenu du choix de codage effectué, choix qui apparaît en tête des sorties, cela signifie qu'il y a augmentation moyenne de 4.402 à la température haute 30° et diminution de cette même quantité 4.402 à la température basse 20°. Par conséquent, en moyenne sur les deux pH, le taux de croissance augmente de $8.804 = 2 \times 4.402$ entre la température basse 20° et la température haute 30°. Mais cette moyenne cache une différence importante entre les situations aux deux pH. Si $pH = -1$ (7.5 avec le choix de codage effectué), le taux de croissance augmente en plus de 3.198 si $T = 1$ tandis qu'il diminue de cette même quantité 3.198 si $T = -1$. Dans ce cas, l'effet température est donc amplifié. Inversement si $pH = 1$ (8.5), il faut retrancher 3.198 quand $T = 1$, l'additionner quand $T = -1$ et l'effet température est nettement amoindri.

Cette gymnastique est un peu délicate et implique de ne pas se tromper sur la correspondance entre niveaux réels et niveaux codés. Une autre façon d'interpréter est de demander l'impression des moyennes pour T et pour le couple $T.\,pH$. C'est ce qui a été fait dans l'analyse figurant au tableau 4. Dans ce cas d'un plan factoriel complet avec un nombre constant de répétitions de chaque traitement, les moyennes figurant dans les sorties sont identiques à celle qu'on obtient directement à partir des données qui apparaissent en haut du tableau 4. Ceci n'est pas vrai dans le cas général.

Dans le cas général, le calcul des moyennes prend en compte le modèle. Tout se passe comme si on calculait la réponse calculée pour chaque combinaison de niveaux des facteurs introduits, puis on faisait les moyennes appropriées de ces réponses. Comme le modèle retenu n'inclue pas nécessairement toutes les interactions entre facteurs, ces réponses calculées peuvent différer des moyennes calculées directement à partir des données.

En fait, l'algorithme de calcul procède différemment. Il n'utilise que les termes du modèle contenant exclusivement les facteurs pour lesquels on demande la moyenne. Ainsi pour calculer les moyennes par niveau de T, on utilise seulement les effets factoriels estimés $\hat{e}(\mathbf{1})$ et $\hat{e}(T)$. La formule donnant la moyenne $\hat{\mu}(T)$ au niveau T est

$$\hat{\mu}(T) = \hat{e}(\mathbf{1}) + T\hat{e}(T) \ .$$

On vérifie qu'elle redonne bien les deux moyennes figurant dans le tableau 4

$$\begin{array}{lcccccc} \hat{\mu}(-1) & = & 5.461 & - & 4.402 & = & 1.059 \\ \hat{\mu}(\ \ 1) & = & 5.461 & + & 4.402 & = & 9.863 \end{array}$$

Les moyennes par niveau du couple T, pH ne sont autres que les estimations $\hat{\tau}(T, pH)$ des réponses théoriques $\tau(T, pH)$. On les obtient en remplaçant dans (6) les effets factoriels

```
DONNEES (fichier EX2P2R.DAT)

  T  pH #tau             y
 30 8.5    8  5.0243570176
 30 8.5    8  5.2990213340
 30 7.5   14 15.2316257890
 30 7.5   14 13.8977458285
 20 8.5    4  5.1478851378
 20 8.5    4  0.3590072791
 20 7.5    2  1.1865591781
 20 7.5    2 -2.4581348916

PARAMETRAGE DE L'ANALYSE

Nom d'analyse : EX2P2R
Plan : c:/exanalys/Ex2p2r.dat
Variables : c:/exanalys/Ex2p2r.dat
         ...
facteurs et covariables :     libelle nat type niveaux
                              T       num fac
                              pH      num fac

              Modele    :    T.pH

             Moyennes   : [1]   T
                          (2]   T.pH

PARAMETRISATION. CONTRASTES POUR

 facteur T                      facteur pH
 niv c0 c1                      niv c0 c1
  20  1 -1                      7.5  1 -1
  30  1  1                      8.5  1  1

ANALYSE de VARIANCE de y

VARIANCE expliquée  84.98   ddl = 3
Variance d'erreur    4.76   ddl = 4

      ddl   F Sned   Prob (%) Signif.
        1   50.1      0.2      **
 pH     1    3.8     12.3
 T      1   32.6      0.5      **
 T.pH   1   17.2      1.4      *

EFFET FACTORIELS

variable  tau                       variable  y
       effet               effet   +/-   +/-    +/-
                                  (95%) (99%) (99.9%)
         7                  5.461 2.141 3.551   6.641
 T       4             T    4.402 2.141 3.551   6.641
 T.pH   -2             T.pH -3.198 2.141 3.551   6.641
 pH     -1             pH   -1.503 2.141 3.551   6.641

MOYENNES

Moyenne sur  T      Moyenne sur le(s) facteur(s)  T pH
    tau     y                tau      y
 20   3 1.059         20 7.5   2 -0.636
 30  11 9.863         20 8.5   4  2.754
                      30 7.5  14 14.564
                      30 8.5   8  5.162
```

TAB. 4 – *Analyse du plan factoriel complet 2^2 avec 2 répétitions (fichier EX2P2R.DAT)* par leurs estimation

$$\hat{\tau}(T.\,pH) = \hat{e}(\mathbf{1}) + T\hat{e}(T) + pH\,\hat{e}(pH) + T\,pH\,\hat{e}(T.\,pH)$$

ce qui redonne également les résultats du tableau 4

$$\begin{array}{lcccccccccc}
\hat{\tau}(-1,-1) & = & 5.461 & - & 4.402 & + & 1.503 & - & 3.198 & = & -0.636, \\
\hat{\tau}(-1,\ \ 1) & = & 5.461 & - & 4.402 & - & 1.503 & + & 3.198 & = & 2.754, \\
\hat{\tau}(\ \ 1,-1) & = & 5.461 & + & 4.402 & + & 1.503 & + & 3.198 & = & 14.564, \\
\hat{\tau}(\ \ 1,\ \ 1) & = & 5.461 & + & 4.402 & - & 1.503 & - & 3.198 & = & 5.162\,.
\end{array}$$

1.6 La fraction 2^{2-1}

Si on ne peut expérimenter que deux traitements sur les $4 = 2^2$ du plan factoriel complet *température*$\times$ *pH* considéré précédemmment, on doit sélectionner une fraction 1/2 de ce plan. Le tableau 5 fournit un choix possible. Les traitements retenus sont

	1	A	B	AB	$AB=-1$	
$\tau(-1,-1)$	1	-1	-1	1		
$\tau(-1,\ 1)$	1	-1	1	-1	$\star$	$\longrightarrow y(-1,\ 1)$
$\tau(\ 1,-1)$	1	1	-1	-1	$\star$	$\longrightarrow y(\ 1,-1)$
$\tau(\ 1,\ 1)$	1	1	1	1		
	$e(\mathbf{1})$	$e(A)$	$e(B)$	$e(AB)$	**effets factoriels**	

TAB. 5 – *Sélection d'une fraction* 2^{2-1}

désignés par une étoile. Ce sont ceux qui vérifie la relation algébrique

$$AB = -1\,. \tag{15}$$

Cette relation est appelée *relation de définition* de la fraction.

Une fraction de taille aussi réduite n'a pas d'intérêt pratique. Elle n'est introduite ici que pour faciliter la présentation. Les fractions définies par une ou plusieurs relations de définition de la même nature que (15) sont dites *régulières.* On a déjà noté en introduction leur intérêt.

Notation: Classiquement, une fraction $1/2^m$ d'un plan factoriel complet pour n facteurs à 2 niveaux est notée 2^{n-m}. Comme 2^n est le nombre de traitement du plan factoriel complet sans répétition, $2^{n-m} = 2^n/2^m$ est le nombre de traitements expérimentés dans la fraction $1/2^m$. Bien sur dans cette notation, il est important de faire apparaître les deux exposants n et m.

La fraction considérée ici est donc notée 2^{2-1}. En multipliant l'égalité (15) par A, on obtient $A^2B = -A$. Comme $A^2 = 1$, ceci conduit à l'égalité $B = -A$.

Les niveaux codés vérifient donc, sur les unités de la fraction

$$AB = -1, \qquad B = -A$$

et le modèle (6)

$$\tau(A{,}B) = e(\mathbf{1}) + Ae(A) + Be(B) + ABe(AB)$$

peut donc s'écrire sous la forme

$$\tau(A{,}B) = [e(\mathbf{1}) - e(AB)] + \quad A\,[e(A) - e(B)] \ . \tag{16}$$

Si on note $\gamma(\mathbf{1})$, $\gamma(A)$ les deux combinaisons linéaires d'effets factoriels y apparaissant:

$$\gamma(\mathbf{1}) = e(\mathbf{1}) - e(AB), \quad \gamma(A) = e(A) - e(B) \ , \tag{17}$$

l'expression (16) devient

$$\tau(A{,}B) = \gamma(\mathbf{1}) + \quad A\gamma(A) \ .$$

Elle conduit aux égalités

$$\begin{array}{rcccc} \tau(-1{,}1) & = & \gamma(\mathbf{1}) & - & \gamma(A) \\ \tau(1,-1) & = & \gamma(\mathbf{1}) & + & \gamma(A) \end{array}$$

qui s'écrivent sous la forme matricielle suivante

$$\begin{array}{ccccc} \begin{bmatrix} \tau(-1,\ \ 1) \\ \tau(\ \ 1,-1) \end{bmatrix} & = & \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} & \times & \begin{bmatrix} \gamma(\mathbf{1}) \\ \gamma(A) \end{bmatrix} \\ \tau_0 & = & Z_0 & & \gamma \ . \end{array}$$

Le vecteur τ_0 à gauche a pour coordonnées les réponses théoriques associées aux deux traitements de la fraction. Le vecteur γ à droite a pour coordonnées les combinaisons linéaires $\gamma(\mathbf{1})$, $\gamma(A)$ de (17). La matrice Z_0 à droite de l'égalité a pour première colonne, associée à $\gamma(\mathbf{1})$, le vecteur $\mathbf{1}$ de coordonnées toutes égales à 1, et pour seconde colonne, associée à $\gamma(A)$ le vecteur A donnant les niveaux de A sur la fraction.

Le modèle linéaire associé s'écrit

$$\begin{array}{ccccccc} \begin{bmatrix} y(-1,\ \ 1) \\ y(\ \ 1,-1) \end{bmatrix} & = & \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} & \times & \begin{bmatrix} \gamma(\mathbf{1}) \\ \gamma(A) \end{bmatrix} & + & \begin{bmatrix} \varepsilon(-1,\ \ 1) \\ \varepsilon(\ \ 1,-1) \end{bmatrix} \\ y & = & Z_0 & & \gamma & + & \varepsilon_0 \end{array}$$

où y est le vecteurs des 2 réponses observées et ε_0 celui des erreurs correspondantes.

On voit que les réponses théoriques et observées ne dépendent des effets factoriels qu'au travers des combinaisons linéaires $\gamma(\mathbf{1})$ et $\gamma(A)$. Ces combinaisons linéaires sont appelées *fonctions estimables de base* ce que l'on abrège en *f.e.b.*.

Si $e(\mathbf{1})$ et $e(AB)$ sont augmentés d'une même quantité, la f.e.b $\gamma(\mathbf{1}) = e(\mathbf{1}) - e(AB)$ et donc les réponses restent inchangées. A partir des deux réponses observées, on peut estimer $\gamma(\mathbf{1}) = e(\mathbf{1}) - e(AB)$ mais on ne peut pas estimer séparément ces deux effets factoriels $e(\mathbf{1})$ et $e(AB)$. On dit qu'ils sont *confondus* par la fraction. De même, on peut estimer $\gamma(A) = e(A) - e(B)$ mais il est impossible à partir des seules données de la fraction d'obtenir des estimations séparées des effets factoriels $e(A)$ et $e(B)$. L'effet $e(A)$ *est confondu* avec $e(B)$ sur la fraction.

On vérifie immédiatement que la matrice Z_0 vérifie

$$Z_0'Z_0 = Z_0Z_0' = 2\,\mathbf{I}\ .$$

On en déduit l'égalité vectorielle

$$\gamma = \frac{1}{2}Z_0'\tau_0$$

et en utilisant les résultats classique du modèle linéaire

$$\hat\gamma = \frac{1}{2}Z_0'y\ .$$

Les coordonnées du vecteur à gauche sont les estimations des f.e.b $\gamma(\mathbf{1})$, $\gamma(A)$ qui sont donc données par

$$\begin{array}{rclclcl}
\hat\gamma(\mathbf{1}) & = & \mathbf{1}'y/\,2 & = & \langle y,\mathbf{1}\rangle/\,2 & = & [\ y(-1,1) + y(1,-1)]/\,2 \\
\hat\gamma(A) & = & A'y/\,2 & = & \langle y,A\rangle/\,2 & = & [-y(-1,1) + y(1,-1)]/\,2
\end{array}$$

Si on suppose les observations indépendantes et qu'on fait l'hypothèse d'homoscédasticité (1) qui s'écrit ici

$$\mathrm{var}(y(A,B)) = \sigma^2$$

on a

$$\mathrm{var}(\hat\gamma) = \frac{\sigma^2}{2}\mathbf{I}_2$$

soit

$$\begin{array}{rcl}
\mathrm{var}\,(\hat\gamma(\mathbf{1})) & = & \sigma^2/2\ , \\
\mathrm{var}\,(\hat\gamma(A)) & = & \sigma^2/2\ , \\
\mathrm{cov}\,(\hat\gamma(\mathbf{1}),\hat\gamma(A)) & = & 0\ .
\end{array}$$

Les estimateurs de f.e.b $\hat\gamma(\mathbf{1})$, $\hat\gamma(A)$ ont donc une covariance nulle. On dit qu'ils sont estimés *orthogonalement.* On trouve par ailleurs qu'ils sont estimés avec la même variance qui est σ^2/N où $N = 2$ est le nombre d'observations de la fraction. Nous retrouverons ces propriétés pour les autres fractions régulières.

On peut entrevoir sur cet exemple l'intérêt de ce type de fraction. Si l'interaction $e(AB)$ est négligée, la f.e.b $\gamma(\mathbf{1}) = e(\mathbf{1}) - e(AB)$ coincide avec la moyenne générale $e(\mathbf{1})$ qui peut alors être estimée. Dans les fractions examinées par la suite, on s'arrange pour confondre chaque effet important uniquement avec des effets négligeables. Cet effet important peut alors être assimilé à la f.e.b qui le contient et peut donc être estimé. Par exemple on ne confondra les effets principaux qu'avec des interactions de 3 facteurs ou plus. On sait que ces dernières sont souvent négligeables.

Facteurs de base et facteurs définis

Pour construire le plan défini par la relation (15), on peut lister l'ensemble des traitements du plan factoriel complet et sélectionner ceux qui vérifie la relation comme cela est fait dans le tableau 5.

Mais il est plus simple de lister tous les niveaux du facteur A et d'en déduire le niveau de B par la relation $B = -A$. On dit alors qu'on a pris A comme *facteur de base* et que B est un *facteur défini* (on dit aussi *ajouté*). Le tableau 6 montre ce mode de formation On notera dans ce cas de figure que l'on peut construire la fraction aussi bien en prenant

	A	B	$= -A$
Construction :	-1	1	
	1	-1	

TAB. 6 – *Construction de la fraction* 2^{2-1} *avec* A *pour facteur de base*

B comme facteur de base et en en déduisant A par la relation $A = -B$.

1.7 Plan fractionnaire régulier 2^{3-1}

On désire maintenant étudier l'influence sur une réponse y de trois facteurs, A, B, C à 2 niveaux. Ces niveaux sont codés 1, -1. Mais seuls 4 traitements peuvent être expérimentés.

La fraction retenue est celle définie par la relation de définition

$$ABC = -1 \ . \tag{18}$$

Le tableau 7 donne les éléments pour passer des réponses théoriques, à gauche, aux effets factoriels, en bas, et indique par une étoile les 4 traitements retenus dans la fraction définie par (18).

	$\mathbf{1}$	A	B	C	AB	AC	BC	ABC	$ABC=-1$	
$\tau(-1,-1,\ 1)$	1	-1	-1	1	1	-1	-1	1		
$\tau(-1,-1,-1)$	1	-1	-1	-1	1	1	1	-1	$\star$	$\longrightarrow y(-1,-1,-1)$
$\tau(-1,\ 1,\ 1)$	1	-1	1	1	-1	-1	1	-1	$\star$	$\longrightarrow y(-1,\ 1,\ 1)$
$\tau(-1,\ 1,-1)$	1	-1	1	-1	-1	1	-1	1		
$\tau(\ 1,-1,\ 1)$	1	1	-1	1	-1	1	-1	-1	$\star$	$\longrightarrow y(\ 1,-1,\ 1)$
$\tau(\ 1,-1,-1)$	1	1	-1	-1	-1	-1	1	1		
$\tau(\ 1,\ 1,\ 1)$	1	1	1	1	1	1	1	1		
$\tau(\ 1,\ 1,-1)$	1	1	1	-1	1	-1	-1	-1	$\star$	$\longrightarrow y(\ 1,\ 1,-1)$
	$e(\mathbf{1})$	$e(A)$	$e(B)$	$e(C)$	$e(AB)$	$e(AC)$	$e(BC)$	$e(ABC)$		
	effets factoriels									

TAB. 7 – *Effets traitement et factoriels pour le plan* 2^3*. Fraction* 2^{3-1}

Comme $A^2 = 1$, $B^2 = 1$, on obtient en multipliant l'égalité (18) successivement par 1, A, B, AB les égalités

$$ABC = -1, \quad BC = -A, \quad AC = -B, \quad C = -AB \tag{19}$$

qui sont vérifiées sur la fraction.

Le modèle exprimant les réponses théoriques en fonction des effets factoriels est de la même forme que (6):

$$\tau(A,B,C) = e(\mathbf{1}) + A\,e(A) + B\,e(B) + C\,e(C) + BC\,e(BC) + AC\,e(AC) + AB\,e(AB) + ABC\,e(ABC)\ .$$

Compte tenu des égalités (19), il se met pour les traitements de la fraction sous la forme

$$\tau(A,B,C) = \underbrace{[e(\mathbf{1}) - e(ABC)]}_{\gamma(\mathbf{1})} + A\,\underbrace{[e(A) - e(BC)]}_{\gamma(A)} + B\,\underbrace{[e(B) - e(AC)]}_{\gamma(B)} + AB\,\underbrace{[e(AB) - e(C)]}_{\gamma(AB)}$$

soit

$$\tau(\ A,\ B,\ C) = \gamma(\mathbf{1}) + A\ \gamma(A) + B\ \gamma(B) + AB\ \gamma(AB)$$

qui conduit aux 4 égalités

$$\begin{array}{rcl} \tau(-1,-1,-1) &=& \gamma(\mathbf{1}) - \gamma(A) - \gamma(B) - \gamma(AB) \\ \tau(-1,\ 1,\ 1) &=& \gamma(\mathbf{1}) - \gamma(A) + \gamma(B) + \gamma(AB) \\ \tau(\ 1,-1,\ 1) &=& \gamma(\mathbf{1}) + \gamma(A) - \gamma(B) + \gamma(AB) \\ \tau(\ 1,\ 1,-1) &=& \gamma(\mathbf{1}) + \gamma(A) + \gamma(B) - \gamma(AB)\ . \end{array}$$

La forme matricielle est

$$\underbrace{\begin{bmatrix} \tau(-1,-1,-1) \\ \tau(-1,\ 1,\ 1) \\ \tau(\ 1,-1,\ 1) \\ \tau(\ 1,\ 1,-1) \end{bmatrix}}_{\tau_0} = \underbrace{\begin{bmatrix} 1 & -1 & -1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & 1 & 1 & 1 \end{bmatrix}}_{Z} \times \underbrace{\begin{bmatrix} \gamma(\mathbf{1}) \\ \gamma(A) \\ \gamma(B) \\ \gamma(AB) \end{bmatrix}}_{\gamma}$$

Elle conduit au modèle linéaire

$$\underbrace{\begin{bmatrix} y(-1,-1,-1) \\ y(-1,\ 1,\ 1) \\ y(\ 1,-1,\ 1) \\ y(\ 1,\ 1,-1) \end{bmatrix}}_{y} = \underbrace{\begin{bmatrix} 1 & -1 & -1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & 1 & 1 & 1 \end{bmatrix}}_{Z} \times \underbrace{\begin{bmatrix} \gamma(\mathbf{1}) \\ \gamma(A) \\ \gamma(B) \\ \gamma(AB) \end{bmatrix}}_{\gamma} + \underbrace{\begin{bmatrix} \varepsilon(-1,-1,-1) \\ \varepsilon(-1,\ 1,\ 1) \\ \varepsilon(\ 1,-1,\ 1) \\ \varepsilon(\ 1,\ 1,-1) \end{bmatrix}}_{\varepsilon\ .}$$

On notera que les colonnes de Z associées à $\gamma(\mathbf{1})$, $\gamma(A)$, $\gamma(B)$, $\gamma(AB)$ sont les vecteurs correspondant aux quantités figurant dans les parenthèses qui suivent γ. La première colonne est le vecteur $\mathbf{1}$ formé de 1, les deux suivantes donnent les niveaux A et B respectivement et la dernière le produit AB. Comme on a sur la fraction $\mathbf{1} = -ABC$, $A = -BC$, $B = -AC$, $AB = -C$ ces colonnes coincident également avec les vecteurs $-ABC$, $-BC$, $-AC$, $-C$.

La matrice Z vérifie

$$Z'Z = ZZ' = 4\,\mathbf{I}$$

et on a donc

$$\gamma = \frac{1}{4} Z' \tau_0 \qquad\qquad \hat{\gamma} = \frac{1}{4} Z' y$$

ce qui s'écrit aussi

$$\begin{aligned}
\hat{\gamma}(\mathbf{1}) &= \mathbf{1}'y/4 = \langle y,\mathbf{1}\rangle/4 = [\, y(-1,-1,-1) + y(-1,1,1) + y(1,-1,1) + y(1,1,-1)]/4\\
\hat{\gamma}(A) &= A'y/4 = \langle y,A\rangle/4 = [\, -y(-1,-1,-1) - y(-1,1,1) + y(1,-1,1) + y(1,1,-1)]/4\\
\hat{\gamma}(B) &= B'y/4 = \langle y,B\rangle/4 = [\, -y(-1,-1,-1) + y(-1,1,1) - y(1,-1,1) + y(1,1,-1)]/4\\
\hat{\gamma}(AB) &= (AB)'y/4 = \langle y,AB\rangle/4 = [\, y(-1,-1,-1) - y(-1,1,1) - y(1,-1,1) + y(1,1,-1)]/4
\end{aligned}$$

Les fonctions estimables de bases (f.e.b) sont

$$\begin{array}{ll}
\gamma(\mathbf{1}) = e(\mathbf{1}) - e(ABC), & \gamma(A) = e(A) - e(BC),\\
\gamma(B) = e(B) - e(AC), & \gamma(AB) = e(AB) - e(C)\ .
\end{array}$$

En supposant l'indépendance entre les 4 observations effectuées et l'homoscédasticité

$$\mathrm{var}\big(y(A,B,C)\big) = \sigma^2$$

on obtient

$$\mathrm{var}(\hat{\gamma}) = \frac{\sigma}{4}\mathbf{I}_4$$

soit

$$\begin{array}{rclrcl}
\mathrm{var}\,(\hat{\gamma}(\mathbf{1})) &=& \sigma^2/4, & \mathrm{var}\,(\hat{\gamma}(A)) &=& \sigma^2/4,\\
\mathrm{var}\,(\hat{\gamma}(B)) &=& \sigma^2/4, & \mathrm{var}\,(\hat{\gamma}(AB)) &=& \sigma^2/4,\\
\mathrm{cov}\,(\hat{\gamma}(\mathbf{1}),\hat{\gamma}(A)) &=& 0, & \mathrm{cov}\,(\hat{\gamma}(\mathbf{1}),\hat{\gamma}(B)) &=& 0,\\
\mathrm{cov}\,(\hat{\gamma}(\mathbf{1}),\hat{\gamma}(AB)) &=& 0, & \mathrm{cov}\,(\hat{\gamma}(A),\hat{\gamma}(B)) &=& 0,\\
\mathrm{cov}\,(\hat{\gamma}(A),\hat{\gamma}(AB)) &=& 0, & \mathrm{cov}\,(\hat{\gamma}(B),\hat{\gamma}(AB)) &=& 0.
\end{array}$$

En d'autre termes $\gamma(\mathbf{1})$, $\gamma(A)$, $\gamma(B)$, $\gamma(AB)$ sont estimés orthogonalement avec une même variance σ^2/N où $N = 4$ est le nombre d'observations.

Les f.e.b peuvent être estimées. Mais il est impossible d'obtenir des estimations séparées des effets factoriels apparaissant dans chacune d'elles. On dit qu'ils sont confondus. Donc dans cette fraction

$$\begin{array}{lll@{\qquad}lll}
e(\mathbf{1}) & \text{est confondu avec} & e(ABC), & e(A) & \text{est confondu avec} & e(BC)\\
e(B) & \text{est confondu avec} & e(AC), & e(AB) & \text{est confondu avec} & e(C)
\end{array}$$

Mais si on néglige les interactions, on peut estimer la moyenne générale et les effets principaux. En terme statistique

$$\begin{array}{ll}
e(\mathbf{1}) \text{ est estimable si } e(ABC) = 0, & e(A) \text{ est estimable si } e(BC) = 0\\
e(B) \text{ est estimable si } e(AC) = 0, & e(C) \text{ est estimable si } e(AB) = 0
\end{array}$$

Facteurs de base et facteurs définis

Comme dans le cas de la fraction 2^{2-1}, la fraction définie par (18) peut être construite en listant les combinaisons de niveaux des facteurs A, B et en déduisant le niveau du

	A	B	C	$= -AB$
	-1	-1	1	
Construction :	-1	1	-1	
	1	-1	-1	
	1	1	1	

TAB. 8 – *Construction de la fraction* 2^{3-1} *avec* A*,* B *pour facteur de base*

facteur C par l'égalité $C = -AB$. On dit alors que A et B sont pris comme facteurs de base et que C est un facteur défini. Le tableau tableau 8 montre ce mode de formation

Dans ce cas, on peut également prendre comme facteurs de base soit le couple (A, C), soit le couple (B, C).

1.8 Fraction 2^{6-2}

Dans ce nouvel exemple, on a 6 facteurs A, B, C, D, E, F à 2 niveaux codés 1, -1 et 16 unités expérimentales. Contrairement aux plans précédents, un peu maigre pour répondre a des besoins pratiques, le plan proposé ici peut être utilisé dans de nombreuses circonstances.

Ce plan est construit en prenant A, B, C, D comme *facteur de base* et en utilisant les relations de définition

$$E = ABC, \qquad F = -BCD \tag{20}$$

pour trouver les valeurs des *facteurs définis* E et F.

Il faut noter dans ce cas que l'on ne peut pas prendre n'importe quel sous-ensemble de 4 facteurs comme facteurs de base. Par exemple il est impossible de prendre A, B, C, E comme facteurs de base car l'égalité $E = ABC$ interdit d'avoir dans le plan toutes les 16 combinaisons de niveaux des 4 facteurs A, B, C, E.

On a $ABC = E \implies ABCE = E^2 \implies ABCE = 1$ et de même $-BCDF = 1$. En multipliant les deux dernières égalités, on obtient $-ABCEBCDF = 1$, soit $AB^2C^2DEF = 1$ ou encore $ADEF = 1$. Les deux relations de définition (20) conduisent donc aux égalités

$$\mathbf{1} \quad = \quad ABCE \quad = \quad -BCDF \quad = \quad -ADEF \tag{21}$$

Les produits de facteurs apparaissant dans ces égalités sont appelés *mots, produits ou contrastes de définition.* La relation $ADEF = -1$ déduite des deux autres relations est également appelée une *relation de définition.* S'il y a lieu, on peut préciser pour les relations figurant dans (20) qu'il s'agit des *relations de définition génératrices.*

En multipliant les égalités (21) successivement par les 2^4 produits que l'on peut former avec les facteurs de base, on obtient 16 systèmes d'égalités

$$\begin{array}{ccccccc}
\mathbf{1} & = & ABCE & = & -BCDF & = & -ADEF \\
A & = & BCE & = & -ABCDF & = & -DEF \\
B & = & ACE & = & -CDF & = & -ABDEF \\
AB & = & CE & = & -ACDF & = & -BDEF \\
C & = & ABE & = & -BDF & = & -ACDEF \\
AC & = & BE & = & -ABDF & = & -CDEF \\
BC & = & AE & = & -DF & = & -ABCDEF \\
ABC & = & E & = & -ADF & = & -BCDEF \\
D & = & ABCDE & = & -BCF & = & -AEF \\
AD & = & BCDE & = & -ABCF & = & -EF \\
BD & = & ACDE & = & -CF & = & -ABEF \\
ABD & = & CDE & = & -ACF & = & -BEF \\
CD & = & ABDE & = & -BF & = & -ACEF \\
ACD & = & BDE & = & -ABF & = & -CEF \\
BCD & = & ADE & = & -F & = & -ABCEF \\
ABCD & = & DE & = & -AF & = & -BCEF
\end{array} \tag{22}$$

De même que dans l'exemple de la fraction 2^{3-1}, ces égalités donnent directement les effets confondus et les f.e.b qui sont

$$\begin{array}{lllllll}
\gamma(\mathbf{1}) & = e(\mathbf{1}) & + & e(ABCE) - & e(BCDF) & - & e(ADEF), \\
\gamma(A) & = e(A) & + & e(BCE) - & e(ABCDF) - & & e(DEF), \\
\gamma(AB) & = e(AB) & + & e(CE) - & e(ACDF) & - & e(BDEF), \\
 & & & \dots & & & \\
\gamma(ABCD) & = e(ABCD) + & & e(DE) - & e(AF) & - & e(BCEF).
\end{array}$$

Si on néglige les interactions de trois facteurs ou plus, les f.e.b deviennent

$$\begin{array}{ll}
\gamma(\mathbf{1}) & = e(\mathbf{1}) \\
\gamma(A) & = e(A) \\
\gamma(AB) & = e(AB) + e(CE) \\
 & \dots \\
\gamma(ABCD) & = e(DE) - e(AF)
\end{array}$$

Les formules pour l'estimation et pour les variances, covariances sont analogues à celles obtenues pour la fraction 2^{3-1}:

$$\begin{array}{lcllcl}
\hat{\gamma}(\mathbf{1}) & = & \langle y,\mathbf{1}\rangle / 16 & \mathrm{var}\,(\hat{\gamma}(\mathbf{1})) & = & \sigma^2/16 \\
 & \dots & & \mathrm{var}\,(\hat{\gamma}(A)) & = & \sigma^2/16 \\
\hat{\gamma}(A) & = & \langle y,A\rangle / 16 & \mathrm{var}\,(\hat{\gamma}(AB)) & = & \sigma^2/16 \\
\hat{\gamma}(AB) & = & \langle y,AB\rangle / 16 & \mathrm{cov}\,(\hat{\gamma}(\mathbf{1}),\hat{\gamma}(AB)) & = & 0 \\
 & \dots & & \dots & &
\end{array}$$

où A, B, AB, ... sont les vecteurs de dimension 16 donnant les niveaux A, B, les produits AB, ... pour chacune des observations.

Plan en notation multiplicative et additive

On a jusqu'à présent utilisé des niveaux codés 1 et -1 et défini les facteurs ajoutés, au signe près, comme des produits. Mais pour programmer, il existe un autre codage plus pratique. Les niveaux sont codés 0 et 1, au lieu de $1 = (-1)^0$ et $-1 = (-1)^1$. Les produits sont remplacés par des sommes modulo 2 et s'ils sont précédés d'un signe $-$, on ajoute 1 à cette somme avant de prendre le reste de la division par 2 pour calculer le résultat modulo 2. Le tableau 9 montre le plan avec les deux codages. On parle de notation additive quand on utilise le codage en 0, 1 et des sommes modulo 2, de notation multiplicative quand les niveaux sont codés 1 et -1. La notation additive peut être assez facilement généralisée pour obtenir des fractions de plan quand tous les facteurs ont 3 niveaux (fraction 3^{n-m}). L'addition modulo 2 est alors remplacée par l'addition modulo 3.

$E = A + B + C \pmod 2$
$F = 1 + B + C + D \pmod 2$

A	B	C	D	E	F
0	0	0	0	0	1
0	0	0	1	0	0
0	0	1	0	1	0
0	0	1	1	1	1
0	1	0	0	1	0
0	1	0	1	1	1
0	1	1	0	0	1
0	1	1	1	0	0
1	0	0	0	1	1
1	0	0	1	1	0
1	0	1	0	0	0
1	0	1	1	0	1
1	1	0	0	0	0
1	1	0	1	0	1
1	1	1	0	1	1
1	1	1	1	1	0

notation additive

$E = ABC$
$F = -BCD$

A	B	C	D	E	F
1	1	1	1	1	−1
1	1	1	−1	1	1
1	1	−1	1	−1	1
1	1	−1	−1	−1	−1
1	−1	1	1	−1	1
1	−1	1	−1	−1	−1
1	−1	−1	1	1	−1
1	−1	−1	−1	1	1
−1	1	1	1	−1	−1
−1	1	1	−1	−1	1
−1	1	−1	1	1	1
−1	1	−1	−1	1	−1
−1	−1	1	1	1	1
−1	−1	1	−1	1	−1
−1	−1	−1	1	−1	−1
−1	−1	−1	−1	−1	1

notation multiplicative

TAB. 9 – *Fraction* $1/4$ *d'un* 2^6

Confusion et longueur des mots de définition

Pour trouver les effets confondus avec $e(A)$, on multiplie les produits dans (21) par A. La multiplication par $\mathbf{1}$ redonne A, la multiplication par les 3 autres produits donne les effets confondus avec A : $e(BCE)$, $e(ABCDF)$, $e(DEF)$. On remarque que le produit par A soit fait disparaître ce facteur du produit (cas de $A \times ABCE$, $A \times ADEF$), soit l'y rajoute – cas de $A \times BCDF$. Dans tous les cas, comme les 3 produits de définition de (21) contiennent 4 facteurs, la multiplication par A laisse subsister au moins 3 facteurs. Donc $e(A)$ n'est confondu qu'avec des interactions de 3 facteurs au moins. Il clair qu'il en est de même pour les autres effets principaux $e(B), \ldots, e(F)$. Parce que les mots de définition comprennent 4 facteurs ou plus, ils ne sont confondus qu'avec des interactions d'au moins 3 facteurs.

De même, pour trouver les effets factoriels confondus avec une interaction telle $e(AB)$, on multiplie les mots dans (21) par AB. Ceci fait disparaître au plus deux lettres. Comme les 3 produits de définition ont 4 lettres, $e(AB)$ n'est confondu qu'avec des interactions de 2 facteurs ou plus. Il en est de même des autres interactions de deux facteurs. Elles peuvent être confondues avec d'autres interactions de deux facteurs, mais pas avec des effets principaux.

Y a-t-il des interactions de 2 facteurs qui ne se confondent qu'avec des interactions de

3 facteurs ou plus, donc qui soient estimables sous l'hypothèse de nullité des interactions de 3 facteurs? La réponse est non. On peut le vérifier en examinant les égalités (22), ou par un examen direct de l'occurrence des $15 = C_6^2$ paires de deux lettres apparaissant dans les mots de définition (tableau 10). Chaque paire apparaît dans un mot au moins

mot déf.	couples de lettres dans les mots de définition														
$ABCE$	AB	CE	AC	BE	AE	BC									
$BCDF$						BC	DF	BD	CF	BF	CD				
$ADEF$					AE		DF					AD	EF	AF	DE

TAB. 10 – *Paires de lettres dans les mots de définition*

Il y a 6 paires de lettres dans chaque mot de définition, mais certaines paires apparaissent dans deux mots

et l'interaction correspondante est donc confondue avec au moins une autre interaction. Les paires AE, BC, DF apparaissant dans deux mots conduisent à trois interactions confondues entre elles.

Ce petit exemple montre que l'inspection directe des mots de définition, et en particulier de leurs longueurs, donne déjà des indications importantes sur les propriétés du plan. La longueur minimum de ces mots, égale à 4 dans la fraction définie par (21), est appelée la *résolution* du plan.

Definition 1 (Fraction régulière de résolution k) *Une fraction régulière 2^{n-p} est dite de résolution r si r est la longueur minimum des $2^p - 1$ mots de définition.*

Cette notion très importante est étudiée plus en détail dans le paragraphe qui suit.

1.9 Plans réguliers de résolution 3, 4, 5.

Nous ne considérons ici que des fractions *régulières* qui sont les fractions définies par des relations algébriques comme dans les exemples précédents.

1.9.1 Résolution 3

Considérons la fraction 2^{5-2} ayant 3 facteurs de base A, B, C et deux facteurs supplémentaires D, E définis par

$$D = AB, \qquad E = AC \,. \tag{23}$$

Il est immédiat que les niveaux des 5 facteurs sur cette fraction vérifient

$$\mathbf{1} \quad = \quad ABD \quad = \quad ACE \quad = \quad BCDE \tag{24}$$

Ces égalités donne les 3 mots de définition, ABD, ACE de longueur 3 et $BCDE$ de longueur 4. La longueur minimale est 3 et le plan est par définition de résolution 3. La multiplication par un facteur fait disparaître au plus une lettre et en laisse subsister au moins deux. Donc un effet principal ne peut être confondu qu'avec des interactions, pas avec un autre effet principal. Par exemple, la multiplication par A donne BD, CE, $ABCDE$. Si on néglige l'interaction $e(ABCDE)$ des 5 facteurs, on voit que A est confondu avec deux interactions de deux facteurs $e(BD)$, $e(CE)$, mais pas avec aucun des autres effets principaux $e(B), \ldots, e(F)$. La f.e.b correspondante est

$$\gamma(A) = e(A) + e(BD) + e(CE) \tag{25}$$

Cette propriété importante est la propriété fondamentale des plans de de résolution 3.

Propriété 1 (Propriété des plans de résolution 3) *Dans un plan de résolution* 3, *un effet principal n'est confondu avec aucun autre effet principal. Sous l'hypothèse qu'il n'y a pas d'interaction, tous les effets principaux sont donc estimables*

L'hypothèse d'absence d'interaction est très forte et sans doute rarement vérifiée. Quel peut donc être l'intérêt des plans de résolution 3?

On verra au paragraphe 1.11 qu'ils permettent d'étudier un grand nombre de facteurs avec peu de données. Ceci est intéressant quand on croit que seuls quelques facteurs sont réellement actifs au sein d'un ensemble beaucoup plus important de facteurs susceptibles d'agir. On réalise alors un premier plan de criblage, souvent de résolution 3, pour sélectionner ces facteurs actifs que l'on peut alors inclure dans des études plus spécialisées. Même si ces facteurs interagissent, on peut penser que leurs effets principaux sont très dominants, donc détectables dans un plan de résolution 3. Si l'analyse du plan donnent de nombreuses f.e.b significatives, il faut généralement poursuivre l'expérimentation pour déterminer si ce sont réellement des effets principaux qui sont à l'origine de cette significativité, ou bien des interactions. Par contre, si il y a peu de f.e.b significatives, l'examen détaillé des résultats peut donner des indications assez précises, comme l'illustre l'exercice suivant.

Exercice 1 *Préciser toutes les f.e.b de la fraction définie par (23). Si y désigne le vecteur des 8 réponses observées, préciser les estimateurs de ces f.e.b et leurs variances? Déterminer ensuite ces estimations numériquement pour chacun des deux vecteurs de réponse y_1, y_2 figurant dans le tableau 11. Retrouver les résultats avec ANALYS. Quelles conclusions amène ces résultats pour chacune des deux observations?*

A	B	C	$D = AB$	$E = AC$	y_1	y_2
−1	−1	−1	1	1	33	43
−1	−1	1	1	−1	67	57
−1	1	−1	−1	1	127	37
−1	1	1	−1	−1	173	63
1	−1	−1	−1	−1	37	107
1	−1	1	−1	1	63	193
1	1	−1	1	−1	123	93
1	1	1	1	1	177	207

TAB. 11 – *Plan et données pour une fraction* 2^{5-2}

Les valeurs de y_1, y_2 ont été choisies pour faciliter les calculs

Deux f.e.b ressortent nettement dans l'analyse de y_1, $\gamma(B) = e(B) + e(AD)$, $\gamma(C) = e(C) + e(AE)$. Comme les f.e.b ou apparaissent les effets principaux des facteurs A, D, E sont négligeables, il est très vraisemblable que l'ont peut négliger $e(AD)$, $e(AE)$ devant $e(B)$, $e(C)$ et attribuer aux effets principaux de B et C l'importance de $\gamma(B)$ et $\gamma(C)$.

Dans l'analyse de y_2, trois f.e.b ressortent $\gamma(A) = e(A)+e(BD)$, $\gamma(C) = e(C)+e(AE)$, $\gamma(AC) = e(AC)+e(E)$. Compte tenu de l'importance des deux premières, on ne peut pas exclure que ce soit l'interaction AC qui rende la dernière f.e.b importante et non l'effet de E. Il n'est pas exclu non plus que ce soit au contraire A et E les facteurs les plus actifs et que l'importance de $\gamma(C)$ priovienne de l'interaction AE plutôt que de l'effet de C.

Cet exemple montre qu'un plan de résolution 3 donne d'autant plus d'indications que le nombre de facteurs actif est plus réduit. Aussi dans les cas où les facteurs étudiés ont tous de bonnes chances d'être actifs et d'interagir, un plan de résolution 3 est insuffisant. Une alternative dans ce cas est d'avoir directement recours aux plans de résolution 5 et 4 décrits aux paragraphes suivants.

1.9.2 Résolution 5

Un exemple de plan de résolution 5 est la fraction 2^{8-2} ayant 6 facteurs de base A, B, C, D, E, F et deux facteurs G, H supplémentaires définis par

$$G = ABCD, \qquad H = CDEF\,. \tag{26}$$

Les 3 mots de définition figurent dans (27), avec sous chaque mot sa longueur.

$$\begin{array}{ccccccc} \mathbf{1} & = & ABCDG & = & CDEFH & = & ABGEFH \\ & & 5 & & 5 & & 6 \end{array} \tag{27}$$

La longueur minimum est 5 et cette fraction est donc par définition de résolution 5. Lorsqu'on multiplie par une lettre les mots de définition, il subsiste au moins 4 lettres. Lorsqu'on multiplie par un produit de deux lettres, il subsiste au moins trois lettres. La conséquence est la propriété suivante, valide bien sûr pour n'importe quelle fraction de résolution 5.

Propriété 2 (Propriété des plans de résolution 5) *Dans un plan de résolution* 5*, un effet principal n'est confondu qu'avec des interactions d'au moins* 4 *facteurs. Les interactions de deux facteurs ne sont confondues qu'avec des interactions d'au moins trois facteurs.*

Une conséquence est que dans une telle fraction, on peut estimer les effets principaux et interactions de deux facteurs si on néglige les interactions de 3 facteurs ou plus. On a en fait une propriété beaucoup plus forte. Même si il y a des interactions de 3 facteurs, les effets principaux sont correctement estimés puisque les f.e.b où ils apparaissent n'incluent que des interactions de 4 facteurs ou plus. Ceci montre l'extrême robustesse de ces plans.

Mais nous verrons au paragraphe 1.11 que ces fractions de résolution 5 sont très gourmandes en données. En d'autre termes, elle ne permettent d'étudier qu'un nombre assez limité de facteurs. C'est pourquoi on a souvent recours à des plans de résolution 4 moins gourmands en données, mais qui conduisent néanmoins à de bonnes estimations des effets principaux même s'il y a des interactions.

1.9.3 Résolution 4

Une des propriétés fondamentale de la fraction 2^{6-2} définie par (20) est qu'un effet principal n'y est confondu qu'avec des interactions de 3 facteurs ou plus. Ceci vient de ce que le plan est de résolution 4. Un autre exemple où les mots de définition n'ont pas tous la même longueur est la fraction 2^{7-2} ayant 5 facteurs de base A, B, C, D, E et deux facteurs F, G additionnels définis par

$$F = ABCD, \qquad G = CDE\,. \tag{28}$$

Les 3 mots de définition figurent dans les égalités suivantes, avec sous chaque mot sa longueur.

$$\begin{array}{ccccccc} \mathbf{1} & = & ABCDF & = & CDEG & = & ABEFG \\ & & 5 & & 4 & & 5 \end{array} \tag{29}$$

La longueur minimum est 4 et la fraction est donc par définition de résolution 4. Lorsqu'on multiplie par une lettre les mots de définition, il subsiste au moins 3 lettres. Donc on a la propriété suivante, valable dans toutes les fractions de résolution 4.

Propriété 3 (Propriété des plans de résolution 4) *Dans un plan de résolution* 4, *un effet principal n'est confondu qu'avec des interactions d'au moins* 3 *facteurs.*

En revanche, les interactions de 2 facteurs peuvent être confondues entre elles. La résolution 4 ne garantit donc pas que l'on puisse estimer les interactions de deux facteurs. Mais elle permet d'estimer correctement les effets principaux même si il y a des interactions de deux facteurs.

Ainsi en multipliant (29) successivement par C et CD, on obtient les égalités

$$\begin{array}{ccccccc} C & = & ABDF & = & DEG & = & ABCEFG, \\ CD & = & ABF & = & EG & = & ABCDEFG. \end{array}$$

Les f.e.b correspondantes sont

$$\begin{array}{ccccccccc} \gamma(C) & = & e(C) & + & e(ABDF) & + & e(DEG) & + & e(ABCEFG), \\ \gamma(CD) & = & e(CD) & + & e(ABF) & + & e(EG) & + & e(ABCDEFG). \end{array}$$

Si on néglige les interactions de 3 facteurs et plus, elles s'écrivent

$$\begin{array}{ccc} \gamma(C) & = & e(C), \\ \gamma(CD) & = & e(CD) + e(EG). \end{array}$$

On vérifie donc dans ce cas particulier que, sous l'hypothèse de nullité des interactions de 3 facteurs et plus, l'effet principal de C est estimable même si on ne fait aucune hypothèse sur les interactions de deux facteurs. Les autres effets principaux sont estimables dans les mêmes conditions.

Si l'analyse donne des effets principaux de C et D nettement prédominants, et que la f.e.b $\gamma(CD)$ apparaît significativement différente de 0, il est naturel d'attribuer cette significativité plutôt à l'interaction $e(CD)$ des deux facteurs prédominants qu'à $e(EG)$. Un plan de résolution 4 peut donc dans nombre de cas apporter des informations assez précises sur les interactions même s'il ne permet pas a priori de les estimer individuellement.

1.10 Equivalence entre fractions

Considérons la fraction 2^{5-2} définie par

$$D = ABC, \qquad E = AB\,. \tag{30}$$

Les égalités suivantes donnent ses 3 mots de définition et leur longueurs

$$\begin{array}{ccccccc} \mathbf{1} & = & ABCD & = & ABE & = & CDE \\ & & 4 & & 3 & & 3 \end{array} \tag{31}$$

Cette fraction diffère de celle définie par (23), mais lui ressemble beaucoup. Elle est également de résolution 3, avec 2 mots de définition de longueur 3 et 1 mot de longueur 4. En réalité, il est possible dans ce cas de passer d'une fraction à l'autre par une des permutations des 5 lettres définies par les flèches ci-dessous:

$$\begin{array}{ccc} A & BD & CE \\ \downarrow & \downarrow & \downarrow \\ E & AB & CD \end{array}$$

Une telle permutation envoie A sur E, le groupe de lettre $\{B,D\}$ sur le groupe $\{A,D\}$ et le groupe $\{C,E\}$ sur le groupe $\{C,D\}$. Elle transforme donc les mots de définition $BCDE$, ABD et ACE figurant dans (24) en les mots $ABCD$, ABE, CDE de (31).

On dit de deux telles fractions qu'elles sont *équivalentes* ou encore de *même type*. Pour l'utilisateur, elles sont exactement similaires et interchangables, du moins tant que les lettres ne sont pas encore précisément affectées aux facteurs réels. Car quelle que soit l'affectation effectuée pour l'une des fractions, on peut trouver une affectation donnant strictement le même plan pour l'autre.

Cette dernière observation en appelle une autre. Le rôle des lettres dans ces deux fractions n'est pas symétrique. Si on prend la première fraction (23) par exemple, la lettre A y apparaît dans les deux mots de définition de 3 lettres et pas dans celui de 4 lettres contrairement aux autres lettres. La conséquence est que l'effet principal de A se confond avec deux interactions de deux facteurs BD, CE alors que les autres effets principaux ne se confondent qu'avec une seule interaction de deux facteurs. Si l'on met en place une telle fraction, cette dissymétrie suggère d'affecter A au facteur réel a priori le moins important. Bien sur, si on part de l'autre fraction (30) ou c'est E qui joue le rôle de A, c'est E qui sera affecté au facteur le moins important.

On peut aller plus loin dans l'examen de la fraction (30) en remarquant qu'elle est a priori plus informative sur une interaction telle que BC que sur une interaction comme BD. La première BC, dont les deux lettres n'apparaissent ensemble que dans le mot de définition de 4 lettres, n'est confondue qu'avec une autre interaction DE, tandis que la seconde, présente au sein d'un mot de 3 lettres, est confondue avec un effet principal. S'il y a deux facteurs dont on pense avant l'expérience qu'ils sont importants et susceptibles d'interagir, c'est donc à deux lettres comme B et C qu'on les affectera, plutôt qu'à B et D.

1.11 Majoration du nombre n de facteurs en fonction de la résolution

Pour sélectionner une fraction, il est bon d'avoir une idée du nombre maximum de facteurs que l'on peut introduire avec un nombre donné d'unités expérimentales et une résolution donnée. Nous donnons ici cette limite dans le cas des fractions régulières pour facteurs à 2 niveaux. Le nombre d'unité est dans ce cas une puissance de 2.

1.11.1 Majoration du nombre de facteurs en résolution 3

Avec 2^s unités expérimentales, on peut étudier jusqu'à $2^s - 1$ facteurs.

Le cas $s = 3$ permet d'illustrer clairement la façon de procéder. La règle énoncée donne une limite de $7 = 2^3 - 1$ facteurs. Pour construire la fraction avec A, B, C comme facteurs de base, les facteurs supplémentaires D, E, F, G sont définis par les 4 produits AB, AC, BC, ABC comme indiqué dans le tableau 12.

A	B	C	D $= AB$	E $= AC$	F $= BC$	G $= ABC$
−1	−1	−1	1	1	1	−1
−1	−1	1	1	−1	−1	1
−1	1	−1	−1	1	−1	1
−1	1	1	−1	−1	1	−1
1	−1	−1	−1	−1	1	1
1	−1	1	−1	1	−1	−1
1	1	−1	1	−1	−1	−1
1	1	1	1	1	1	1

TAB. 12 – *Fraction à 8 unités, résolution 3 avec le nb. maxi. 7 de facteurs*

On peut naturellement changer le signe de certains produits, par exemple utiliser $-AB$ au lieu de AB. Cela ne fait qu'intervertir 1 et -1 et n'a donc pas de conséquence si la correspondance entre niveaux réels et niveaux codés n'a pas été précisée.

Exercice 2 *Préciser les mots de définition de la fraction du tableau 12. En déduire les interactions de deux facteurs confondues avec l'effet principal A et donner les f.e.b $\gamma(A)$; $\gamma(B)$ pour cette fraction et la fraction qu'on obtient en remplaçant AB par $-AB$.*

Les fractions des tableaux 11 et 12 ont même nombre d'unités expérimentales et même résolution. En fait, la première s'obtient en éliminant les colonnes F et G de la seconde.

Cette procédure d'élimination est classiquement utilisée. Quels que soient les facteurs éliminés, le plan après élimination a au moins la même résolution que le plan d'où on est parti.

Il y a cependant deux erreurs à ne pas faire. La première est de penser qu'à nombre d'unité et résolution égale, c'est le plan permettant d'étudier le maximum de facteurs qui est obligatoirement le plus intéressant. Ainsi la fraction à 7 facteurs du tableau 12 confond chaque effet principal avec 3 interactions, alors que celle du tableau 11 confond l'effet principal de A avec deux interactions de deux facteurs, les autres effets principaux avec seulement une interaction de deux facteurs. L'interprétation sera donc généralement plus aisée avec cette seconde fraction.

La seconde erreur est de croire que l'on peut choisir totalement au hasard le sous-ensemble de facteurs utilisé lorsqu'on expérimente avec moins de facteurs que le nombre maximal. Dans le cas de la fraction 2^{5-2}, il s'avère que toutes les fractions de résolution 3 sont équivalentes à une permutation des facteurs près (l'équivalence est définie plus précisément au paragraphe 1.10). Donc tous les choix de 5 facteurs parmi les 7 du tableau 12 se valent. Mais si on a à retenir seulement 4 facteurs, les choix ne sont plus équivalent.

Exercice 3 *Pourquoi est-il plus judicieux d'utiliser la fraction 2^{4-1} définie par $D = ABC$ que celle définie par $D = AB$?*

1.11.2 Majoration du nombre de facteurs en résolution 4

Avec 2^s unités expérimentales, on peut étudier jusqu'à 2^{s-1} facteurs en résolution 4. Une façon de former le plan est d'utiliser tous les produits d'un nombre impair de facteurs de base. Par exemple, si $s = 4$ et qu'on prend A, B, C, D comme facteurs de base, on peut définir 4 facteurs supplémentaires E, F, G, H par

$$E = ABC, \qquad F = ABD, \qquad G = ACD, \qquad H = BCD\,. \tag{32}$$

Le plan ainsi formé a le nombre maximum $8 = 2^{4-1}$ de facteurs.

Exercice 4 *Prouver que cette fraction est de résolution 4. Quels sont les mots de définition de longueur 4? En examinant les couples de lettres qui apparaissent dans ces mots, préciser la taille des ensembles d'interactions de deux facteurs confondues entre elles. Préciser la f.e.b qui contient l'interaction AB. Si on suppose nulles les interactions de 3 facteurs ou plus, l'analyse d'un tel plan dégage-t-elle des degrés de liberté pour estimer la variance résiduelle?*

On peut extraire de ce plan des plans de résolution 4 pour 7 et 6 facteurs. En fait le choix des 6 ou 7 facteurs retenus n'a dans ce cas pas d'importance car toutes les fractions régulières 2^{6-2} de résolution 4 sont équivalentes et de même toutes les fractions 2^{7-3} de résolution 4 sont aussi équivalentes.

Soient A, B, C, D, E, F les facteurs d'une fraction régulière 2^{6-2} de résolution 4. On montre qu'on peut toujours la construire à partir de 4 facteurs de base, et quitte à modifier l'appellation des facteurs, on peut supposer que ces 4 facteurs sont A, B, C, D. Les facteurs E et F doivent alors être définis comme des produits d'au moins 3 des 4 facteurs de base, sinon on a un mot de définition de 3 lettres.

Mais si $E = ABCD$, aucun produit de 4 ou 3 lettres ne convient pour F. Ainsi si $F = ABCD$, on a $EF = 1$ donc un mot de définition de 2 lettres. Si F est produit de 3 lettres, on trouve en formant le produit EF un mot de définition de 3 lettres. Par exemple si $F = ABC$, on a $EF = D$, soit $DEF = 1$.

Donc E ne peut être qu'un produit de 3 lettres et F de même un produit de 3 lettres différent de E. Il est clair qu'il n'y a, à une permutation des facteurs A, B, C, D près, qu'un choix possible pour E et F.

Le cas d'une fraction 2^{7-3} de résolution 4 est similaire. Pour les mêmes raisons, les facteurs définis sont des produits distincts de 3 des facteurs de base, et à une permutation près, il n'y a qu'un choix possible.

Si l'on veut expérimenter seulement 5 facteurs avec 16 unités, la sélection ne peut plus être effectuée de façon arbitraire. On a à l'évidence intérêt à utiliser la fraction définie par $E = ABCD$ de résolution 5 de préférence à toute autre. Cela montre que le procédé consistant à extraire certains des facteurs d'une fraction plus importante n'est en règle générale pas satisfaisant.

1.11.3 De la résolution 3 à la résolution 4

Lorsqu'un premier plan de résolution 3 donne trop de f.e.b importantes pour permettre une interprétation convenable, il est toujours possible de le compléter par un plan de même taille de façon à obtenir globalement un plan de résolution 4 qui permet de déterminer avec beaucoup moins de risques quels sont les facteurs réellement actifs.

La procédure consiste simplement à reprendre le même plan en changeant tous les signes. Pour l'analyse, on prend généralement en compte un facteur supplémentaire prenant le niveau 1 sur la première partie du plan, -1 sur la seconde partie. Ce facteur est noté S dans ce qui suit. Son effet donne le décalage moyen entre les résultats des deux parties de plan. Dans la terminologie habituelle des plans d'expériences, les deux parties du plan sont appelées des *blocs* et un tel décalage s'appelle un *effet bloc*. Il s'avère que le plan est encore de résolution 4 avec ce facteur supplémentaire S. Mais cette propriété n'est généralement pas exploitée car on fait généralement l'hypothèse que, même s'il y a un décalage entre les deux parties du plan, l'effet des facteurs y reste similaire. Autrement dit, on suppose qu'il n'y a pas d'interaction entre ces facteurs et le facteur bloc S.

Considérons par exemple la fraction définie par (23). Le tableau 13 redonne à gauche cette fraction et à droite la fraction opposée, avec le facteur supplémentaire S égal à 1 à gauche, -1 à droite. On a rajouté une réponse y inventée ici pour permettre des calculs qui tombent juste et une interprétation simple. Comme tous les signes sont changés à

unités 1 à 8 (plan à 5 fact. de réso. 3							unité 9 à 16 : fraction opposée						
S	A	B	C	D	E	y	S	A	B	C	D	E	y
1	−1	−1	−1	1	1	63	−1	1	1	1	−1	−1	267
1	−1	−1	1	1	−1	67	−1	1	1	−1	−1	1	163
1	−1	1	−1	−1	1	77	−1	1	−1	1	1	−1	93
1	−1	1	1	−1	−1	93	−1	1	−1	−1	1	1	−23
1	1	−1	−1	−1	−1	17	−1	−1	1	1	1	1	53
1	1	−1	1	−1	1	113	−1	−1	1	−1	1	−1	57
1	1	1	−1	1	−1	183	−1	−1	−1	1	−1	1	47
1	1	1	1	1	1	307	−1	−1	−1	−1	−1	−1	23

TAB. 13 – *Doublement par l'opposé d'une fraction de résolution* 3

droite, l'égalité $D = AB$ valable à gauche donne sur la partie droite la relation $(-D) = (-A)(-B)$, soit $D = -AB$. On peut donc écrire à gauche comme à droite $D = SAB$. De même la seconde relation de définition $E = AC$ de la fraction (23) conduit à la relation globale $E = SAC$. En fin de compte la fraction globale peut être définie à partir des facteurs de base A, B, C, S par les relations

$$D = ABS, \qquad E = ACS\ . \tag{33}$$

Il est immédiat de vérifier à partir de ces relations que le plan global à 16 observations a bien la résolution 4: les deux relations (33) conduisent à la troisième relation $BCDE = 1$ qui implique bien 4 facteurs.

Plus généralement, dans ce doublement par l'opposé d'un plan de résolution 3, on voit que

- chaque relation initiale impliquant un nombre pair de facteurs reste vraie sur l'opposé donc globalement
- chaque relation initiale impliquant un nombre impair de facteurs conduit à une relation avec le facteur S en plus.

En particulier, le doublement par l'opposé transforme les mots de définition à 3 lettres en mots à 4 lettres et conserve les mots de définition à 4 lettres.

Exercice 5 *Le tableau 14 donne à gauche les résultats d'analyse de la fraction* 2^{5-3} *à gauche du tableau 13, à droite l'analyse sur les* 16 *résultats obtenus après doublement par l'opposé. Expliquer à la lumière des résultats du plan doublé les résultats obtenus pour la première fraction de résolution* 3.

Analyse variable y							
sur fraction 2^{5-2} (8 unités à gauche du tab 13)				sur fraction doublée par l'opposé (16 unités du tab 13)			
effet		+/− (95%)	+/− (99%)	effet		+/− (95%)	+/− (99%)
	115	16.4	37.8		100	16.4	37.8
B	50	16.4	37.8	B	50	16.4	37.8
A	40	16.4	37.8	A	40	16.4	37.8
D	40	16.4	37.8	$A.B$	40	16.4	37.8
C	30	16.4	37.8	C	30	16.4	37.8
E	25	16.4	37.8	$A.C$	25	16.4	37.8
				S	15	16.4	37.8
				$A.D$	0	16.4	37.8
				$B.D$	0	16.4	37.8
				D	0	16.4	37.8
				$A.E$	0	16.4	37.8
				E	0	16.4	37.8
				$B.C$	0	16.4	37.8
				$B.E$	0	16.4	37.8

TAB. 14 – *Analyse fraction rés.* 3, *puis fraction doublée rés.* 4

1.11.4 Majoration du nombre de facteurs en résolution 5

Il n'y a pas de formule simple donnant le nombre maximum de facteurs que l'on peut introduire dans une fraction régulière de résolution 5 en fonction du nombre d'unités. Des études au cas par cas ont cependant permis d'établir le tableau 15 qui suffit dans la pratique courante. On a déjà rencontré une fraction 2^{8-2} de résolution 5 au § 1.9.2. Ce

s	4	5	6	7	8	9
2^s	16	32	64	128	256	512
h	5	6	8	11	17	≥ 23

TAB. 15 – *Nb. maxi. h de facteurs dans fraction régulière taille* 2^s, *résolution 5*

nombre 8 est d'après le tableau le nombre maximum de facteurs que l'on peut introduire avec $64 = 2^6$ unités si l'on veut disposer d'une fraction régulière de résolution 5. On peut facilement retrouver ce résultat dans ce cas par programme ou par une recherche manuelle. La recherche devient beaucoup plus longue avec un nombre d'unités excédant $2^7 = 128$. Avec 512 unités, l'examen de toutes les possibilités devient si long que la limite exacte n'est pas connue avec certitude. On sait simplement qu'on peut introduire 23 facteurs. Il est possible, mais pas certain qu'on puisse en introduire un $24^{\text{ème}}$.

1.12 Choix d'une fraction

La notion de résolution donne un excellent guide pour s'orienter dans le choix d'une fraction. Au démarrage d'une étude, s'il y a un tri à faire parmi les multiples facteurs susceptibles de jouer, un plan de résolution 3 peut apporter rapidement et à moindre frais une première réponse qui permet de sélectionner des facteurs pertinents pour la suite de l'étude.

Lorsqu'on connait les facteurs importants, on s'oriente vers un plan de résolution 4 ou 5. La résolution 5 est souvent difficile à atteindre à cause du nombre d'expériences requises. La résolution 4 fournit dans de nombreux cas un bon premier compromis pour obtenir des résultats fiables sans être obligé d'écarter a priori des facteurs susceptibles d'avoir un rôle non négligeable.

Mais parmi les dispositifs expérimentaux de résolution donnée, il y a souvent de multiples choix possibles. D'autre part, il est des cas où la connaissance que l'on a des facteurs suggère de prendre en compte certaines interactions et d'en ignorer d'autres. Nous allons examiner dans ce paragraphe comment affiner son choix et prendre en compte les éventuelles informations préalables.

1.12.1 Sélection d'une fraction de résolution donnée

Supposons que l'on veut étudier 7 facteurs avec 32 unités expérimentales. La résolution 5 est inatteignable avec un plan régulier (cf §1.11.4). On peut montrer dans ce cas [Kobilinsky, 1997] qu'à une permutation des facteurs près, les seules fractions régulières de résolution 4 sont celles données dans le tableau 16.

	Définition		Mots de définition				Profil
1	$F = ABCDE$	$G = ABC$	**1**	$ABCDEF$	$ABCG$	$DEFG$	$4_2 6_1$
2	$F = ABCD$	$G = CDE$	**1**	$ABCDF$	$CDEG$	$ABEFG$	$4_1 5_2$
3	$F = ABC$	$G = BCD$	**1**	$ABCF$	$BCDG$	$ADFG$	4_3

TAB. 16 – *Fractions 2^{7-2} régulières de résolution* 4

Profil : une notation comme $4_2 6_1$ indique 2 mots de longueur 4 et 1 de longueur 6

A droite de ce tableau, on indique les longueurs des mots de définition, et chaque longueur est indicée par le nombre de mots de cette longueur. Ce *profil de longueurs des mots de définition* est utile pour comparer les fractions entre elles. On a vu que les confusions entre interactions de deux facteurs viennent des mots de définition de 4 lettres.

On peut donc penser que la fraction 2 qui n'a qu'un seul mot de 4 lettres confond moins d'interactions que la fraction 1 qui a deux mots de 4 lettres et que cette dernière est meilleure que la fraction 3 qui a 4 mots de 4 lettres. L'exercice suivant précise la nature des confusions entre interactions de deux facteurs et confirme que la fraction la plus intéressante est la fraction 2 de profil $4_1 5_2$ qui n'a qu'un seul mot de longueur 4.

Exercice 6 *En supposant nulles les interactions de plus de deux facteurs, montrer qu'il y a*

- 6 *groupes de deux interactions confondues et* 9 *interactions non confondues dans la fraction* 1,
- 3 *groupes de deux interactions confondues et* 15 *interactions non confondues dans la fraction* 2,
- 1 *groupe de trois interactions confondues,* 6 *groupes de deux interactions confondues et* 6 *interactions non confondues dans la fraction* 3.

En règle générale, à moins d'avoir des idées précises sur la nature des interactions, on sélectionne plutôt le plan qui permet d'estimer un nombre maximum d'interactions. Il n'y a malheureusement pas d'étude systématique précisant ce plan, mais il est vraisemblable que c'est ce critère du nombre minimal d'interactions confondues qui a servi à sélectionner le catalogue de fractions pour facteurs à 2 niveaux reproduit dans McLean & Anderson (1984).

L'exemple de la fraction 2^{7-2} montre qu'en résolution 4, le nombre d'interactions de deux facteurs confondues entre elles est lié au nombre de mots de définition de 4 lettres. Se basant sur une telle constatation, Fries et Hunter (1980) ont proposé de choisir parmi les plans ayant la résolution r la plus forte possible, celui ou ceux qui ont le nombre minimum de mots de définition de longueur r. Si il en existe plusieurs ayant le même nombre de mots de longueur r, ils proposent celui qui a le moins de mots de longueur $r+1$ et ainsi de suite. Le ou les plans sélectionnés de cette façon sont dit d'*aberration minimale.*

Definition 2 (Fraction d'aberration minimum) *Une fraction 2^{n-p} est dite d'aberration minimum si*

- *sa résolution r est la résolution maximale pour une fraction 2^{n-p},*
- *son nombre W_r de mots de définition de longueur r est minimal parmi les fractions 2^{n-p} de résolution r,*
- *son nombre W_{r+1} de mots de définition de longueur $r+1$ est minimal parmi les fractions de résolution r ayant W_r mots de définition de longueur r, etc ...*

Exercice 7 *Le tableau 16 donne, à permutation près des facteurs, les trois fractions 2^{7-2} de résolution maximum 4. Laquelle est d'aberration minimum?*

On sait trouver une fraction d'aberration minimale 2^{n-p} quand $p \leq 5$, c'est à dire pour des fractions 1/2, 1/4, 1/8, 1/16 et 1/32 de plan 2^n et quel que soit le nombre n de facteurs. Pour des valeurs de p plus grandes, il y a aussi quelques résultats en particulier quand p et $n-p$ sont inférieurs à 9. (Robillard, 1968, Franklin, 1984, Chen et Wu, 1991, Chen, 1992). Ces résultats sont détaillés dans les paragraphes qui suivent.

L'aberration minimale ne garantit pas que la fraction permet d'estimer un nombre maximal d'interactions. On a cependant en résolution 4 le résultat très partiel suivant.

Proposition 1 *Le nombre minimum k_2 d'interactions de 2 facteurs confondues entre elles dans une fraction régulière 2^{n-p} ayant W_4 mots de longueur 4 est donné, pour $W_4 \leq 7$, par le tableau 17. Pour $W_4 > 7$, on a $k_2 \geq 21$.*

W_4	1	2	3	4	5	6	7	>7
k_2	6	12	15	21	24	28	21	≥ 21

TAB. 17 – *Nombre minimum d'interactions confondues en résolution* 4

En utilisant ce résultat, on peut dans certains cas montrer qu'une fraction de résolution 4 et aberration minimale minimise le nombre d'interactions de deux facteurs confondues.

Exercice 8 *Les fractions 2^{7-3}, 2^{8-3}, 2^{9-3} d'aberration minimum sont données au tableau 21. Leurs profils de longueur de mots sont respectivement 4_7, $4_3 5_4$, $4_1 5_4 6_2$. Montrer en utilisant PLANOR qu'elles confondent respectivement* 21, 15, 6 *interactions de deux facteurs. En déduire en utilisant la proposition 1 que ces fractions sont aussi optimales du point de vue du nombre d'interactions de deux facteurs qu'elles permettent d'estimer.*

Exercice 9 *Les fractions 2^{9-4}, 2^{10-4} d'aberration minimum sont données au tableau 26. Leurs profils de longueur de mots sont respectivement $4_6 5_8 8_1$, $4_2 5_8 6_4 8_1$. Montrer en utilisant PLANOR qu'elles confondent respectivement* 28 *et* 12 *interactions de deux facteurs. Ces deux fractions sont-elles optimales du point de vue du nombre d'interactions de deux facteurs qu'elles permettent d'estimer? (on comparera la première à la fraction 2^{9-4} définie par :*

$$F = ABCD, \quad G = ABE, \quad H = ACE, \quad I = BCE\ )$$

Exercice 10 *Le tableau 28 donne les fractions 2^{10-5}, 2^{11-5} et 2^{12-5} d'aberration minimum. Plus précisément, les matrices génératrices des mots de définition de ces fractions sont formées en retenant dans la matrice au bas du tableau les colonnes associées aux* 1 *des lignes $n =$ 10, 11, 12.*

Les profils de longueurs de mots associés sont respectivement $4_{10} 5_{16} 8_5$, $4_4 5_{14} 6_8 8_3 9_2$, $4_1 5_8 6_{12} 7_8 8_1 12_1$.

Montrer que ces 3 *fractions confondent un nombre minimal d'interactions de deux facteurs.*

Note *: pour la première, on montrera avec PLANOR qu'il est impossible en résolution* 4 *d'estimer ne serait ce qu'une seule interaction de deux facteurs.*

1.12.2 Sélection d'une fraction 1/2.

On a clairement intérêt dans ce cas à prendre comme mot de définition le produit de tous les facteurs. Par exemple, si on cherche une fraction 2^{6-1} avec les 6 facteurs A, B, C, D, E, F, on sélectionnera en général la fraction définie par $ABCDEF = 1$ plutôt que celle définie par $ABCDE = 1$. Dans la première de ces fractions, de résolution 6, les interactions de 2 facteurs sont confondues uniquement avec des interactions de 4 facteurs. Elles sont donc estimables même si il y a des interactions de 3 facteurs. En revanche la seconde où n'apparaît pas la lettre F est de résolution seulement 5 et confond les interactions de 2 facteurs avec des interactions de 3 facteurs.

Bien entendu, si un des facteurs est spécialement important et que l'on s'intéresse beaucoup à la façon dont il interagit avec les autres facteurs, la seconde fraction où F n'apparait pas peut être préférable. C'est alors à F qu'est affecté le facteur important.

Exercice 11 *Expliquer pourquoi le plan défini par $ABCDE = 1$ peut être préférable à celui défini par $ABCDEF = 1$ dans le cas où on s'intéresse particulièrement au facteur F et à ses interactions avec les autres facteurs.*

1.12.3 Apparition des facteurs dans les relations de définition

Sauf dans des cas particuliers comme celui de l'exercice, il est préférable d'utiliser une fraction où tous les facteurs apparaissent dans les mots de définition. Si un facteur n'apparait pas, il est en effet clair qu'en le rajoutant à une des relations de définition génératrice, on rallonge tous les mots de définition et on améliore globalement la qualité du plan. On a en particulier

Proposition 2 *Dans une fraction d'aberration minimale, les facteurs doivent tous figurer dans les mots de définition.*

Considérons par exemple la fraction 3 du tableau 16 où le facteur E n'apparait pas. Si on rajoute E dans le produit définissant F, la lettre E est rajoutée à 2 mots de définition, et on obtient une fraction de profil $4_1 5_2$, équivalente à la fraction 2 du tableau, qui n'a plus qu'un mot de longueur 4 et donc seulement 6 interactions de deux facteurs confondues au lieu de 15.

1.12.4 Nombre moyen de lettres par mot de définition

Quand une lettre (i.e. un facteur) apparait dans au moins un mot de définition d'une fraction 2^{n-p}, on montre qu'elle apparait dans exactement la moitié des 2^p mots de définition. Considérons alors une fraction 2^{n-p}, non nécessairement d'aberration minimale, où toutes les n lettres apparaissent. Chacune d'elle apparait donc 2^{p-1} fois dans les mots de définition. Donc le nombre total de lettres figurant dans ces mots est $n2^{p-1}$ et la longueur moyenne d'un mot est $n/2$ si on prend en compte le mot $\mathbf{1}$ de longueur 0 et $n2^{p-1}/(2^p - 1)$ si on ne le prend pas en compte.

Exercice 12 *Utilisant ces formules, donner le nombre moyen de lettres par mot de définition dans une fraction 2^{7-2} si toutes les lettres y apparaissent. Vérifier la formule avec les fractions* 1 *et* 2 *du tableau 16.*

1.12.5 Fraction 1/4 d'aberration minimale

Il y a deux relations de définition génératrices et, compte tenu de ce qui a été dit au paragraphe 1.12.3, chaque facteur doit apparaître dans au moins une relation. Les facteurs peuvent alors être partitionnés en trois groupes selon qu'ils apparaissent dans les deux relations, ou exclusivement dans la première ou la seconde. Si on note P_1, P_2, P_3 les produits des facteurs dans chacun de ces groupes, les relations génératrices s'écrivent

$$P_1P_2 = 1, \quad P_1P_3 = 1 \tag{34}$$

et donne par produit la troisième relation $P_2P_3 = 1$.

Considérons par exemple la fraction 2^{6-2} définie par (20). Le tableau 18 indique les occurences des facteurs (colonnes) dans les mots de définition (lignes). La case associée

mot	A	B	C	D	E	F
$ABCE$	1	1	1	0	1	0
$BCDF$	0	1	1	1	0	1

TAB. 18 – *Matrice génératrice de la fraction* 2^{6-2} *définie par (20)*

à un facteur et un mot contient 1 si le facteur apparaît dans le mot, 0 sinon. La matrice formée de ces 1 et 0 est appelée *matrice génératrice* de la fraction. La partition en trois groupes définie par l'appartenance aux deux relations de définition est clairement indiquée par cette matrice. Les facteurs figurant en tête d'un même vecteur sont dans le même groupe. Les trois groupes de facteurs sont donc $\{A,E\}$, $\{B,C\}$, $\{D,F\}$ qui correspondent respectivement aux vecteurs $(1,0)'$, $(1,1)'$, $(0,1)'$. Si on note $P_2 = AE$, $P_1 = BC$, $P_3 = DF$ les produits associés, (34) redonne la définition (20) au signe près.

Pour que le plan soit d'aberration minimum, il faut simplement former les 3 groupes de *tailles aussi égales que possible.* Nous donnons dans le tableau 19 les fractions obtenues de cette façon avec de 7 à 12 facteurs.

nb. fact	produits P_i	rel. déf.	profil	réso
$n = 7$	$P_1 = ABC$, $P_2 = DE$, $P_3 = GH$	$1 = ABCDE = ABCGH = DEGH$	$4_1 5_2$	4
$n = 8$	$P_1 = ABC$, $P_2 = DEF$, $P_3 = GH$	$1 = ABCDEF = ABCGH = DEFGH$	$5_2 6_1$	5
$n = 9$	$P_1 = ABC$, $P_2 = DEF$, $P_3 = GHI$	$1 = ABCDEF = ABCGHI = DEFGHI$	6_3	6
$n = 10$	$P_1 = ABCJ$, $P_2 = DEF$, $P_3 = GHI$	$1 = ABCJDEF = ABCJGHI = DEFGHI$	$6_1 7_2$	6
$n = 11$	$P_1 = ABCJ$, $P_2 = DEFK$, $P_3 = GHI$	$1 = ABCJDEFK = ABCJGHI = DEFKGHI$	$7_2 8_1$	7
$n = 12$	$P_1 = ABCJ$, $P_2 = DEFK$, $P_3 = GHIL$	$1 = ABCJDEFK = ABCJGHIL = DEFKGHIL$	8_3	8

TAB. 19 – *Fraction* 2^{n-2} *d'aberration minimum pour* $7 \leq n \leq 12$

De façon générale, si la division entière de $2n$ par 3 donne

$$2n = 3q + r \quad \text{où } 0 \leq r < 3,$$

la résolution d'une telle fraction est q et son profil $q_{3-r}(q+1)_r$. Noter que $2n/3$ est la longueur moyenne des mots différents de **1**.

1.12.6 Fraction 1/8 d'aberration minimum

Il y a 3 relations de définition. Les facteurs peuvent donc être divisés en 7 groupes en fonction de leurs occurrences dans ces 3 relations. Comme dans le cas de la fraction 1/4, on divise donc les facteurs en 7 groupes de taille aussi égales que possible. Si la division entière de n par 7 donne

$$n = 7m + s \quad \text{avec le reste } s < 7, \tag{35}$$

il y a s groupes de $m+1$ facteurs et $7-s$ de m facteurs. On note respectivement $P_1, \dots,$ P_s les produits des facteurs dans chacun des s groupes de $m+1$ facteurs, $P_{s+1}, \dots\ P_7$ les produits des facteurs dans chacun des groupes de m facteurs.

Les relations de définition donnant à équivalence près la fraction d'aberration minimum sont

$$P_1P_2P_3P_4 = 1, \quad P_1P_2P_6P_7 = 1, \quad P_1P_3P_5P_7 = 1\ . \tag{36}$$

Ces relations engendrent les égalités suivantes

$$\begin{array}{ccccccccc} 1 & = & P_1P_2P_3P_4 & = & P_1P_2P_6P_7 & = & P_3P_4P_6P_7 & = \\ P_1P_3P_5P_7 & = & P_2P_4P_5P_7 & = & P_2P_3P_5P_6 & = & P_1P_4P_5P_7 \end{array}$$

Le tableau 20 précise le profil de cette fraction en fonction de s et de la partie entière de $4n/7$, longueur moyenne des mots de définition différents de $\mathbf{1}$. Cette partie entière notée q est exprimée en fonction de m et s. Le tableau 21 précise les règles de définition et les profils lorsque $6 \leq n \leq 13$. Les facteurs sont notés par des lettres majuscules et minuscules. Lorsque n dépasse 7 et que l'on est amené à former des groupes de deux facteurs, on note ces deux facteurs par la majuscule et la minuscule d'une même lettre.

n	s	q	profil	résolution
$7m+0$	0	$4m$	q_7	q
$7m+1$	1	$4m$	$q_3 \quad (q+1)_4$	q
$7m+2$	2	$4m+1$	$(q-1)_1 \quad q_4 \quad (q+1)_2$	$q-1$
$7m+3$	3	$4m+1$	$q_3 \quad (q+1)_3\ (q+2)_1$	q
$7m+4$	4	$4m+2$	$q_6 \quad (q+2)_1$	q
$7m+5$	5	$4m+2$	$q_2 \quad (q+1)_4\ (q+2)_1$	q
$7m+6$	6	$4m+3$	$q_4 \quad (q+1)_3$	q

TAB. 20 – *Profil des fractions 2^{n-3} d'aberration minimum*

Les fractions de résolution 4 dans le tableau 21 sont aussi les fractions confondant un nombre minimal d'interactions de deux facteurs (cf exercice 8)

Exercice 13 *La recherche par PLANOR d'une dizaine de fractions 2^{8-3} de résolution 4, avec un ordre aléatoire d'exploration donne les trois profils 4_35_4, 4_56_2, 4_7 avec comme relations de définition possibles pour les deux derniers profils:*

$$\begin{array}{llll} \textit{profil } 4_56_2 & F = BCE, & G = ACE, & H = ABE \\ \textit{profil } 4_7 & F = ACD, & G = ADE, & H = BDE \end{array}$$

n	q	s	P_1	P_2	P_3	P_4	P_5	P_6	P_7	Rel. déf.			profil	réso.
6	3	6	A	B	C	D	E	F		$D = ABC$	$F = AB$	$E = AC$	$3_4\ 4_3$	3
7	4	0	A	B	C	D	E	F	G	$D = ABC$	$F = ABG$	$E = ACG$	4_7	4
8	4	1	Aa	B	C	D	E	F	G	$D = AaBC$	$F = AaBG$	$E = AaCG$	$4_3\ 5_4$	4
9	5	2	Aa	Bb	C	D	E	F	G	$D = AaBbC$	$F = AaBbG$	$E = AaCG$	$4_1\ 5_4\ 6_2$	4
10	5	3	Aa	Bb	Cc	D	E	F	G	$D = AaBbCc$	$F = AaBbG$	$E = AaCcG$	$5_3\ 6_3\ 7_1$	5
11	6	4	Aa	Bb	Cc	Dd	E	F	G	$D = AaBbCcd$	$F = AaBbG$	$E = AaCcG$	$6_6\ 8_1$	6
12	6	5	Aa	Bb	Cc	Dd	Ee	F	G	$D = AaBbCcd$	$F = AaBbG$	$E = AaCcGe$	$6_2\ 7_4\ 8_1$	6
13	7	6	Aa	Bb	Cc	Dd	Ee	Ff	G	$D = AaBbCcd$	$F = AaBbGf$	$E = AaCcGe$	$7_4\ 8_3$	7

TAB. 21 – *Définition et profils de fractions* 2^{n-3} *d'aberration minimum* $(6 \leq n \leq 13)$

Trouver dans chaque cas le nombre d'interactions de deux facteurs confondues. Si l'objectif est d'estimer le plus grand nombre possible de ces interactions, quel est entre les deux profils $4_5 6_2$ *et* 4_7 *celui que vous choisiriez? Quel est le lien entre la fraction de profil* 4_7 *et la fraction* 2^{7-3} *du tableau 21? Dans quel cas cette fraction de profil* 4_7 *peut-elle être préférable à la fraction d'aberration minimum?*

1.12.7 Matrice génératrice et vecteur de fréquences d'une fraction

Les mots de définition sont souvent remplacés, comme au tableau 18, par la *matrice génératrice* qui précise les occurences des facteurs (colonnes) dans les mots de définition (lignes). Par exemple les fractions 2^{7-2} du tableau 16 sont représentées par les matrices à gauche du tableau 22 dont les deux lignes correspondent aux relations de définition génératrices. A droite du même tableau sont données des versions simplifiées de ces même matrices, dans lesquelles les facteurs figurant en haut de colonnes identiques sont regroupés. De tels facteurs sont systématiquement associés dans les mots de définition.

Au dessous de chaque groupe de facteur figure la colonne correspondante de 0 et 1. On a aussi indiqué le nombre de facteurs dans ce groupe. Tant que les lettres n'ont pas reçu de sens concret, la donnée de ces nombres de facteurs suffit pour définir la fraction. La liste de ces nombres est appelée *vecteur de fréquences* de la fraction. En fait c'est ce vecteur de fréquence qui apporte réellement toute l'information. Une façon de préciser une fraction consiste donc à préciser, pour chacun des 2^p vecteurs formables de 0 et 1, sa fréquence c'est à dire le nombre de fois où il apparaît dans la matrice génératrice. A partir de ces fréquences, il est aisé de reconstituer la matrice génératrice, donc les mots de définition.

Le tableau 23 redonne sous forme de vecteurs de fréquences l'information du tableau 21. La matrice figurant au bas du tableau est la matrice génératrice des relations de définition (36). La règle donnée pour avoir l'aberration minimum est que les nombres de facteurs dans les produits P_1 à P_7 associés aux colonnes de cette matrice soit aussi égaux que possible. S'il y a 7 facteurs, il y en a donc obligatoirement 1 par produit et ce n'est qu'à partir de 8 facteurs que l'on voit apparaître des produits contenant 2 facteurs. De même, ce n'est qu'à partir de $15 = 1 + 2 \times 7$ facteurs que l'on verra apparaître un groupe contenant 3 facteurs. Si le nombre n de facteurs est donné par (35), les s premières fréquences doivent donc être égales à $m + 1$ et les $7 - s$ dernières à m.

Exercice 14 *Partant du tableau 23, donner les relations de définition de la fraction* 2^{8-3}

	Matrice génératrice								déf. par les fréquences				
		A	B	C	D	E	F	G		DEF	G	ABC	
1	mot 1	1	1	1	1	1	1	0	nb. fact.	3	1	3	
	mot 2	1	1	1	0	0	0	1	mot 1	1	0	1	
									mot 2	0	1	1	
		A	B	C	D	E	F	G		ABF	EG	CD	
2	mot 1	1	1	1	1	0	1	0	nb. fact.	3	2	2	
	mot 2	0	0	1	1	1	0	1	mot 1	1	0	1	
									mot 2	0	1	1	
		A	B	C	D	E	F	G		E	AF	DG	BC
3	mot 1	1	1	1	0	0	1	0	nb. fact.	1	2	2	2
	mot 2	0	1	1	1	0	0	1	mot 1	0	1	0	1
									mot 2	0	0	1	1

TAB. 22 – *Matrices génératrices des mots de définition et vecteurs de fréquences pour les fraction du tableau 16*

n	fréquences							profil	réso.
6	1	1	1	1	1	1	0	$3_4\ 4_3$	3
7	1	1	1	1	1	1	1	4_7	4
8	2	1	1	1	1	1	1	$4_3\ 5_4$	4
9	2	2	1	1	1	1	1	$4_1\ 5_4\ 6_2$	4
10	2	2	2	1	1	1	1	$5_3\ 6_3\ 7_1$	5
11	2	2	2	2	1	1	1	$6_6\ 8_1$	6
12	2	2	2	2	2	1	1	$6_2\ 7_4\ 8_1$	6
13	2	2	2	2	2	2	1	$7_4\ 8_3$	7
	1	1	1	1	0	0	0		
	1	1	0	0	0	1	1		
	1	0	1	0	1	0	1		
	P_1	P_2	P_3	P_4	P_5	P_6	P_7		
	colonnes associées								

TAB. 23 – *Vecteur de fréquences des fractions* 2^{n-3} *d'aberration minimum pour* $6 \leq n \leq 13$

Pour former la matrice génératrice, répéter la colonne associée à chaque fréquence le nombre de fois indiqué par cette fréquence *d'aberration minimum.*

1.12.8 Fractions 1/16 d'aberration minimum

Il y a 4 relations de définition génératrices. Les facteurs peuvent donc être divisés en 15 groupes en fonction de leurs occurrences dans ces 4 relations. Comme précédemment, on divise les facteurs en 15 groupes de taille aussi égales que possible.

Si la division entière de n par 15 donne

$$n = 15m + s \quad \text{avec le reste } s < 15,$$

il y a s groupes de $m+1$ facteurs et $15-s$ de m facteurs. On note respectivement P_1, ... , P_s les produits des facteurs dans chacun des s groupes de $m+1$ facteurs, P_{s+1}, ... P_{15} les produits des facteurs dans chacun des groupes de m facteurs.

Les relations de définition génératrices de la fraction d'aberration minimum sont données dans le tableau 24. Elles ne sont pas les mêmes dans le cas où $s \neq 5$ et dans le cas $s = 5$. On passe du premier ensemble au second en remplaçant P_{15} par P_5 dans la première relation génératrice. Le tableau 25 précise le profil de cette fraction en fonction

cas $s \neq 5$	$P_1P_6P_7P_8P_9P_{12}P_{14}P_{15} = 1$,	$P_2P_5P_7P_8P_9P_{11}P_{13}P_{15} = 1$,
	$P_3P_5P_6P_8P_{10}P_{11}P_{14}P_{15} = 1$,	$P_4P_5P_6P_7P_{10}P_{12}P_{13}P_{15} = 1$
cas $s = 5$	$P_1P_5P_6P_7P_8P_9P_{12}P_{14} = 1$,	$P_2P_5P_7P_8P_9P_{11}P_{13}P_{15} = 1$,
	$P_3P_5P_6P_8P_{10}P_{11}P_{14}P_{15} = 1$,	$P_4P_5P_6P_7P_{10}P_{12}P_{13}P_{15} = 1$.

TAB. 24 – *Génération des fractions 2^{n-4} d'aberration minimum*

de s et de la partie entière de $8n/15$, longueur moyenne des mots de définition différents de **1**. Cette partie entière notée q est exprimée en fonction de m et s. Le tableau 26 précise les règles de définition et les profils lorsque $5 \leq n \leq 19$. Le tableau 27 redonne cette même information sous forme de vecteurs de fréquences.

Les deux dernières lignes de ce tableau des fréquences 27 ont été mises pour montrer le mécanisme périodique qui permet de former les lignes de fréquences à partir d'un certain nombre de facteurs. On passe de n facteurs à $n+15$ facteurs en ajoutant des 1 à chaque fréquence. Comme la fraction à 15 facteurs pour laquelle toutes les fréquences sont 1 a tous ses 15 mots de définition de longueur 8, les mots de définition de la fraction à $n+15$ facteurs comportent tous 8 facteurs de plus que la fraction correspondante pour n facteurs. L'intérêt pratique de ce mécanisme périodique est cependant faible puisque les fractions 2^{n-4} ont pour $n \geq 20$ au moins 2^{16} unités.

1.12.9 Fractions 1/32 d'aberration minimum

Les vecteurs de fréquences et profil, repris dans Chen J. (1992) sont donnés au tableau 28. Les vecteurs de fréquences pour $n > 44$ facteurs s'obtiennent en rajoutant 1 aux fréquences pour $n - 31$ facteurs. Le tableau se forme donc par un mécanisme périodique mais seulement à partir de ce rang 44. En est-il de même pour les nombres entre 32 et 37 et 41 non donnés dans la table? Cela n'apparaît pas clairement dans l'article de Chen J.

n	s	q	profil				réso.
$15m+0$	0	$8m$	q_{15}				q
$15m+1$	1	$8m$	q_7	$(q+1)_8$			q
$15m+2$	2	$8m+1$	$(q-1)_3$	q_8	$(q+1)_4$		$q-1$
$15m+3$	3	$8m+1$	$(q-1)_1$	q_6	$(q+1)_6$	$(q+2)_2$	$q-1$
$15m+4$	4	$8m+2$	$(q-1)_4$	q_6	$(q+1)_4$	$(q+2)_1$	$q-1$
$15m+5$	5	$8m+2$	q_{10}	$(q+2)_5$			q
$15m+6$	6	$8m+3$	$(q-1)_3$	q_8	$(q+1)_3$	$(q+3)_1$	$q-1$
$15m+7$	7	$8m+3$	q_7	$(q+1)_7$	$(q+4)_1$		q
$15m+8$	8	$8m+4$	q_{14}	$(q+4)_1$			q
$15m+9$	9	$8m+4$	q_6	$(q+1)_8$	$(q+4)_1$		q
$15m+10$	10	$8m+5$	$(q-1)_2$	q_8	$(q+1)_4$	$(q+3)_1$	$q-1$
$15m+11$	11	$8m+5$	q_6	$(q+1)_6$	$(q+2)_2$	$(q+3)_1$	q
$15m+12$	12	$8m+6$	q_{12}	$(q+2)_3$			q
$15m+13$	13	$8m+6$	q_4	$(q+1)_8$	$(q+2)_3$		q
$15m+14$	14	$8m+7$	q_8	$(q+1)_7$			q

TAB. 25 – *Profil des fractions 2^{n-4} d'aberration minimum*

n	q	s	Rel. déf.				profil	réso.
5	2	5	$A=E$	$B=E$	$C=E$	$D=E$	2_{10} 4_5	2
6	3	6	$A=F$	$B=E$	$C=EF$	$D=EF$	2_3 3_8 4_3 6_1	2
7	3	7	$A=FG$	$B=EG$	$C=EF$	$D=EFG$	3_7 4_7 7_1	3
8	4	8	$A=FGH$	$B=EGH$	$C=EFH$	$D=EFG$	4_{14} 8_1	4
9	4	9	$A=FGHI$	$B=EGHI$	$C=EFH$	$D=EFG$	4_6 5_8 8_1	4
10	5	10	$A=FGHI$	$B=EGHI$	$C=EFHJ$	$D=EFGJ$	4_2 5_8 6_4 8_1	4
11	5	11	$A=FGHI$	$B=EGHIK$	$C=EFHJK$	$D=EFGJ$	5_6 6_6 7_2 8_1	5
12	6	12	$A=FGHIL$	$B=EGHIK$	$C=EFHJK$	$D=EFGJL$	6_{12} 8_3	6
13	6	13	$A=FGHIL$	$B=EGHIKM$	$C=EFHJK$	$D=EFGJLM$	6_4 7_8 8_3	6
14	7	14	$A=FGHILN$	$B=EGHIKM$	$C=EFHJKN$	$D=EFGJLM$	7_8 8_7	7
15	8	0	$A=FGHILNO$	$B=EGHIKMO$	$C=EFHJKNO$	$D=EFGJLMO$	8_{15}	8
16	8	1	$A=aFGHILNO$	$B=EGHIKMO$	$C=EFHJKNO$	$D=EFGJLMO$	8_7 9_8	8
17	9	2	$A=aFGHILNO$	$B=bEGHIKMO$	$C=EFHJKNO$	$D=EFGJLMO$	8_3 9_8 10_4	8
18	9	3	$A=aFGHILNO$	$B=bEGHIKMO$	$C=cEFHJKNO$	$D=EFGJLMO$	8_1 9_6 10_6 11_2	8
19	10	4	$A=aFGHILNO$	$B=bEGHIKMO$	$C=cEFHJKNO$	$D=dEFGJLMO$	9_4 10_6 11_4 12_1	9

TAB. 26 – *Définition et profils des fractions 2^{n-4} d'aberration minimum pour $5 \leq n \leq 19$*

n	fréquences															profil	réso.
5	1	1	1	1											1	2_{10} 4_5	2
6	1	1	1	1	1	1										2_3 3_8 4_3 6_1	2
7	1	1	1	1	1	1	1									3_7 4_7 7_1	3
8	1	1	1	1	1	1	1	1								4_{14} 8_1	4
9	1	1	1	1	1	1	1	1	1							4_6 5_8 8_1	4
10	1	1	1	1	1	1	1	1	1	1						4_2 5_8 6_4 8_1	4
11	1	1	1	1	1	1	1	1	1	1	1					5_6 6_6 7_2 8_1	5
12	1	1	1	1	1	1	1	1	1	1	1	1				6_{12} 8_3	6
13	1	1	1	1	1	1	1	1	1	1	1	1	1			6_4 7_8 8_3	6
14	1	1	1	1	1	1	1	1	1	1	1	1	1	1		7_8 8_7	7
15	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	8_{15}	8
16	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	8_7 9_8	8
17	2	2	1	1	1	1	1	1	1	1	1	1	1	1	1	8_3 9_8 10_4	8
18	2	2	2	1	1	1	1	1	1	1	1	1	1	1	1	8_1 9_6 10_6 11_2	8
19	2	2	2	2	1	1	1	1	1	1	1	1	1	1	1	9_4 10_6 11_4 12_1	9
20	2	2	2	2	1	1	1	1	1	1	1	1	1	1	2	10_{10} 12_5	10
21	2	2	2	2	2	2	1	1	1	1	1	1	1	1	1	10_3 11_8 12_3 14_1	10
	1	0	0	0	0	1	1	1	1	0	0	1	0	1	1		
	0	1	0	0	1	0	1	1	1	0	1	0	1	0	1		
	0	0	1	0	1	1	0	1	0	1	1	0	0	1	1		
	0	0	0	1	1	1	1	0	0	1	0	1	1	0	1		
	colonnes associées																

TAB. 27 – *Vecteur de fréquences des fractions* 2^{n-4} *d'aberration minimum pour* $5 \leq n \leq 21$

mais est vraisemblable compte tenu du fait que les premières lignes de la table n'ont sinon aucune raison de figurer.

n	fréquences	profil	réso.
1	1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	$0_{15}1_{16}$	0
2	1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	$0_{7}1_{16}2_{8}$	0
3	1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	$0_{3}1_{12}2_{12}3_{4}$	0
4	1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	$0_{1}1_{8}2_{12}3_{8}4_{2}$	0
5	1 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	$1_{5}2_{10}3_{10}4_{5}5_{1}$	1
6	1 1 1 1 1 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	$2_{15}4_{15}6_{1}$	2
7	1 1 1 1 1 1 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0	$2_{5}3_{12}4_{7}5_{4}6_{3}$	2
8	1 1 1 1 1 0 0 0 0 0 1 0 0 0 1 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0	$2_{1}3_{10}4_{11}5_{4}6_{3}7_{2}$	2
9	1 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 1 1 1 0 1 0 0 0 0 0 0 0 0 0 0	$3_{4}4_{14}5_{8}7_{4}8_{1}$	3
10	1 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 1 1 1 0 1 0 0 0 1 0 0 0 0 0 0	$4_{10}5_{16}8_{5}$	4
11	1 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 1 1 1 0 1 0 0 0 1 0 0 0 0 0 1	$4_{4}5_{14}6_{8}8_{3}9_{2}$	4
12	1 1 1 1 1 1 1 1 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	$4_{1}5_{8}6_{12}7_{8}8_{1}12_{1}$	4
13	1 1 1 1 1 1 1 1 1 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	$5_{3}6_{12}7_{12}8_{3}13_{1}$	5
14	1 1 1 1 1 1 1 1 1 1 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	$6_{7}7_{16}8_{7}14_{1}$	6
15	1 1 1 1 1 1 1 1 1 1 1 1 1] 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	$7_{15}8_{15}15_{1}$	7
16	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	$8_{30}16_{1}$	8
17	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0	$8_{14}9_{16}16_{1}$	8
18	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0	$8_{6}9_{16}10_{8}16_{1}$	8
19	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0	$8_{2}9_{12}10_{12}11_{4}16_{1}$	8
20	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 0 1 0 0 0 0 0 0 0 0 0 0	$9_{8}10_{12}11_{8}12_{2}16_{1}$	9
21	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 0 1 0 0 0 1 0 0 0 0 0 0	$10_{20}12_{10}16_{1}$	10
22	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 0 0 0 0 0 0 0 0 0	$10_{6}11_{16}12_{6}14_{2}16_{1}$	10
23	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 0 0 0 0 0 0 0 0	$11_{14}12_{14}15_{2}16_{1}$	11
24	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 0 0 0 0 0 0 0	$12_{28}16_{3}$	12
25	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 0 0 0 0 0 0	$12_{12}13_{16}16_{3}$	12
26	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 0 0 0 0 0	$12_{4}13_{16}14_{8}16_{3}$	12
27	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 0 0 0 0	$13_{12}14_{12}15_{4}16_{3}$	13
28	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 0 0 0	$14_{24}16_{7}$	14
29	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 0 0	$14_{8}15_{16}16_{7}$	14
30	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1] 1 1 1 1 1 1 1 1 1 1 0	$15_{16}16_{15}$	15
31	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	16_{31}	16
38	1 2 1 1 2 2 1 1 1 2 1 1 1 1 1 1 1 2 1 2 1 1 1 1 1 2 1 2 1 0 1	$18_{4}19_{16}20_{6}22_{4}24_{1}$	18
39	1 2 1 1 2 2 1 1 1 2 1 1 1 1 1 1 1 2 1 2 1 1 1 1 1 2 1 2 2 0 1	$19_{12}20_{14}23_{4}24_{1}$	19
40	1 2 1 1 2 2 1 1 1 2 1 1 1 1 1 1 1 2 1 2 1 1 1 1 1 2 1 2 2 0 2	$20_{26}24_{5}$	20
42	1 2 1 2 2 2 1 1 2 2 2 1 1 1 1 2 1 1 1 1 1 2 1 2 1 2 1 2 1 0 1	$20_{3}21_{16}22_{8}24_{3}28_{1}$	20
43	1 2 1 2 2 2 1 1 2 2 2 1 1 1 1 2 1 1 1 1 1 2 1 2 1 2 1 2 2 0 1	$21_{11}22_{12}23_{4}24_{3}29_{1}$	21
44	1 2 1 2 2 2 1 1 2 2 2 1 1 1 1 2 1 1 1 1 1 2 1 2 1 2 1 2 2 0 2	$22_{23}24_{7}30_{1}$	22
	1 0 0 0 0 1 1 1 1 1 0 1 0 0 0 1 0 1 1 0 1 0 0 1 1 0 0 1 0 1 1		
	0 1 0 0 0 1 1 1 0 0 1 0 1 1 0 1 1 0 1 0 1 0 1 0 1 0 1 0 1 0 1		
	0 0 1 0 0 1 0 0 1 1 0 0 1 1 1 1 1 1 0 0 1 1 0 0 1 1 0 0 1 1 0		
	0 0 0 1 0 0 1 0 1 0 1 1 1 0 1 1 1 1 1 1 0 0 0 0 1 1 1 1 0 0 0		
	0 0 0 0 1 0 0 1 0 1 1 1 0 1 1 1 1 1 1 1 1 1 1 1 0 0 0 0 0 0 0		
	colonnes associées		

TAB. 28 – *Vecteur de fréquences de fractions 2^{n-5} d'aberration minimum*

1.12.10 Fractions $1/2^p$ d'aberration minimum quand $p > 5$

On dispose de résultats donnés par Franklin (1984) pour $p \leq 9$ et $n-p \leq 9$, reproduits dans les tableaux 29 et 30, 31, 32, 33. Seules les fractions marquées par DM, issues d'une recherche systématique menée par Draper et Mitchell (1968), sont garanties d'aberration minimum. Il n'est pas certain que les autres soient d'aberration minimum, mais elles ne doivent pas en être loin et fournissent de bonnes fractions de référence.

n	p		profil
11	6		4_{25} 6_{27} 8_{10} 10_{1}
12	6		4_{6} 5_{24} 6_{16} 8_{9} 9_{8}
13	6	DM	4_{2} 5_{16} 6_{18} 7_{10} 8_{9} 9_{4} 10_{2} 11_{2}
14	6	DM	5_{9} 6_{18} 7_{16} 8_{7} 9_{6} 10_{6} 13_{1}
15	6	DM	6_{25} 8_{30} 10_{3} 12_{5}
13	7		4_{14} 5_{28} 6_{24} 7_{24} 8_{17} 9_{12} 10_{8}
14	7		4_{3} 5_{24} 6_{36} 7_{16} 8_{11} 9_{24} 10_{12} 12_{1}
15	7	DM	5_{15} 6_{30} 7_{26} 8_{15} 9_{16} 10_{18} 11_{6} 13_{1}
16	7	DM	6_{44} 8_{45} 10_{28} 12_{10}
14	8		4_{22} 5_{40} 6_{36} 7_{56} 8_{49} 9_{24} 10_{20} 11_{8}
15	8		5_{7} 6_{32} 7_{52} 8_{40} 9_{35} 10_{48} 11_{28} 12_{8} 13_{5}
16	8		5_{24} 6_{44} 7_{40} 8_{45} 9_{40} 10_{28} 11_{24} 12_{10}
17	8		6_{68} 8_{85} 10_{68} 12_{34}
15	9		4_{30} 5_{60} 6_{60} 7_{105} 8_{105} 9_{60} 10_{60} 11_{30} 15_{1}
16	9		4_{10} 5_{48} 6_{72} 7_{80} 8_{90} 9_{80} 10_{72} 11_{48} 12_{10} 16_{1}
17	9	DM	5_{34} 6_{68} 7_{68} 8_{85} 9_{85} 10_{68} 11_{68} 12_{34} 17_{1}
18	9	DM	6_{102} 8_{153} 10_{153} 12_{102} 18_{1}

TAB. 29 – *Profils de fractions 2^{n-p} d'aberration minimum ou proche du minimum*

rel.déf. $n = 11, \quad p = 6$

	A	B	C	D	E	F	G	H	I	J	K
$A = HIJ$	1	0	0	0	0	0	0	1	1	1	0
$B = IJK$	0	1	0	0	0	0	0	0	1	1	1
$C = GJK$	0	0	1	0	0	0	1	0	0	1	1
$D = GHK$	0	0	0	1	0	0	1	1	0	0	1
$E = GHI$	0	0	0	0	1	0	1	1	1	0	0
$F = GHIJK$	0	0	0	0	0	1	1	1	1	1	1

rel.déf. $n = 12, \quad p = 6$

	A	B	C	D	E	F	G	H	I	J	K	L
$A = HIJL$	1	0	0	0	0	0	0	1	1	1	0	1
$B = GIJK$	0	1	0	0	0	0	1	0	1	1	1	0
$C = HJKL$	0	0	1	0	0	0	0	1	0	1	1	1
$D = GIKL$	0	0	0	1	0	0	1	0	1	0	1	1
$E = GHJL$	0	0	0	0	1	0	1	1	0	1	0	1
$F = GHIK$	0	0	0	0	0	1	1	1	1	0	1	0

rel.déf. $n = 13, \quad p = 6$

	A	B	C	D	E	F	G	H	I	J	K	L	M
$A = IJKL$	1	0	0	0	0	0	0	0	1	1	1	1	0
$B = JKM$	0	1	0	0	0	0	0	0	0	1	1	0	1
$C = HIJM$	0	0	1	0	0	0	0	1	1	1	0	0	1
$D = GHIKLM$	0	0	0	1	0	0	1	1	1	0	1	1	1
$E = GKLM$	0	0	0	0	1	0	1	0	0	0	1	1	1
$F = GHJK$	0	0	0	0	0	1	1	1	0	1	1	0	0

rel.déf. $n = 14, \quad p = 6$

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
$A = GHIJK$	1	0	0	0	0	0	1	1	1	1	1	0	0	0
$B = GHIL$	0	1	0	0	0	0	1	1	1	0	0	1	0	0
$C = GHJLM$	0	0	1	0	0	0	1	1	0	1	0	1	1	0
$D = GHKLMN$	0	0	0	1	0	0	1	1	0	0	1	1	1	1
$E = HIJKLMN$	0	0	0	0	1	0	0	1	1	1	1	1	1	1
$F = GJLN$	0	0	0	0	0	1	1	0	0	1	0	1	0	1

rel.déf. $n = 15, \quad p = 6$

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
$A = GHIJK$	1	0	0	0	0	0	1	1	1	1	1	0	0	0	0
$B = GHILM$	0	1	0	0	0	0	1	1	1	0	0	1	1	0	0
$C = GHINO$	0	0	1	0	0	0	1	1	1	0	0	0	0	1	1
$D = GHJLN$	0	0	0	1	0	0	1	1	0	1	0	1	0	1	0
$E = GIKMO$	0	0	0	0	1	0	1	0	1	0	1	0	1	0	1
$F = GJKLMNO$	0	0	0	0	0	1	1	0	0	1	1	1	1	1	1

TAB. 30 – *Fractions 2^{n-p} d'aberration minimum ou presque, $p = 6$, $n - p \leq 9$*

rel.déf. $n = 13, \quad p = 7$

rel.déf.	A	B	C	D	E	F	G	H	I	J	K	L	M
$A = KLM$	1	0	0	0	0	0	0	0	0	0	1	1	1
$B = JKM$	0	1	0	0	0	0	0	0	0	1	1	0	1
$C = HJK$	0	0	1	0	0	0	0	1	0	1	1	0	0
$D = HIJM$	0	0	0	1	0	0	0	1	1	1	0	0	1
$E = HIKLM$	0	0	0	0	1	0	0	1	1	0	1	1	1
$F = HKL$	0	0	0	0	0	1	0	1	0	0	1	1	0
$G = IJKL$	0	0	0	0	0	0	1	0	1	1	1	1	0

rel.déf. $n = 14, \quad p = 7$

rel.déf.	A	B	C	D	E	F	G	H	I	J	K	L	M	N
$A = HIJ$	1	0	0	0	0	0	0	1	1	1	0	0	0	0
$B = HILN$	0	1	0	0	0	0	0	1	1	0	0	1	0	1
$C = HJLM$	0	0	1	0	0	0	0	1	0	1	0	1	1	0
$D = IJLMN$	0	0	0	1	0	0	0	0	1	1	0	1	1	1
$E = HIKM$	0	0	0	0	1	0	0	1	1	0	1	0	1	0
$F = HJKMN$	0	0	0	0	0	1	0	1	0	1	1	0	1	1
$G = IJKN$	0	0	0	0	0	0	1	0	1	1	1	0	0	1

rel.déf. $n = 15, \quad p = 7$

rel.déf.	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
$A = HIJO$	1	0	0	0	0	0	0	1	1	1	0	0	0	0	1
$B = HILNO$	0	1	0	0	0	0	0	1	1	0	0	1	0	1	1
$C = HJLMO$	0	0	1	0	0	0	0	1	0	1	0	1	1	0	1
$D = IJLMNO$	0	0	0	1	0	0	0	0	1	1	0	1	1	1	1
$E = HIKMO$	0	0	0	0	1	0	0	1	1	0	1	0	1	0	1
$F = HJKMN$	0	0	0	0	0	1	0	1	0	1	1	0	1	1	0
$G = IJKN$	0	0	0	0	0	0	1	0	1	1	1	0	0	1	0

rel.déf. $n = 16, \quad p = 7$

rel.déf.	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
$A = HIJLP$	1	0	0	0	0	0	0	1	1	1	0	1	0	0	0	1
$B = IJKMP$	0	1	0	0	0	0	0	0	1	1	1	0	1	0	0	1
$C = JKLNP$	0	0	1	0	0	0	0	0	0	1	1	1	0	1	0	1
$D = KLMOP$	0	0	0	1	0	0	0	0	0	0	1	1	1	0	1	1
$E = HLMNP$	0	0	0	0	1	0	0	1	0	0	0	1	1	1	0	1
$F = IMNOP$	0	0	0	0	0	1	0	0	1	0	0	0	1	1	1	1
$G = HJNOP$	0	0	0	0	0	0	1	1	0	1	0	0	0	1	1	1

TAB. 31 – *Fractions 2^{n-p} d'aberration minimum ou presque, $p = 7$, $n - p \leq 9$*

rel.déf. $n = 14, \quad p = 8$

rel.déf.	A	B	C	D	E	F	G	H	I	J	K	L	M	N
$A = IJKM$	1	0	0	0	0	0	0	0	1	1	1	0	1	0
$B = IJKN$	0	1	0	0	0	0	0	0	1	1	1	0	0	1
$C = IJKLMN$	0	0	1	0	0	0	0	0	1	1	1	1	1	1
$D = ILM$	0	0	0	1	0	0	0	0	1	0	0	1	1	0
$E = JLN$	0	0	0	0	1	0	0	0	0	1	0	1	0	1
$F = JMN$	0	0	0	0	0	1	0	0	0	1	0	0	1	1
$G = KLN$	0	0	0	0	0	0	1	0	0	0	1	1	0	1
$H = KMN$	0	0	0	0	0	0	0	1	0	0	1	0	1	1

rel.déf. $n = 15, \quad p = 8$

rel.déf.	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
$A = IJKMN$	1	0	0	0	0	0	0	0	1	1	1	0	1	1	0
$B = IJLMO$	0	1	0	0	0	0	0	0	1	1	0	1	1	0	1
$C = IKLNO$	0	0	1	0	0	0	0	0	1	0	1	1	0	1	1
$D = MNO$	0	0	0	1	0	0	0	0	0	0	0	0	1	1	1
$E = JKL$	0	0	0	0	1	0	0	0	0	1	1	1	0	0	0
$F = KLMN$	0	0	0	0	0	1	0	0	0	0	1	1	1	1	0
$G = JLNO$	0	0	0	0	0	0	1	0	0	1	0	1	0	1	1
$H = IJKLM$	0	0	0	0	0	0	0	1	1	1	1	1	1	0	0

rel.déf. $n = 16, \quad p = 8$

rel.déf.	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
$A = IJKM$	1	0	0	0	0	0	0	0	1	1	1	0	1	0	0	0
$B = JKLN$	0	1	0	0	0	0	0	0	0	1	1	1	0	1	0	0
$C = KLMO$	0	0	1	0	0	0	0	0	0	0	1	1	1	0	1	0
$D = LMNP$	0	0	0	1	0	0	0	0	0	0	0	1	1	1	0	1
$E = IMNO$	0	0	0	0	1	0	0	0	1	0	0	0	1	1	1	0
$F = JNOP$	0	0	0	0	0	1	0	0	0	1	0	0	0	1	1	1
$G = IKOP$	0	0	0	0	0	0	1	0	1	0	1	0	0	0	1	1
$H = IJLP$	0	0	0	0	0	0	0	1	1	1	0	1	0	0	0	1

rel.déf. $n = 17, \quad p = 8$

rel.déf.	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
$A = IJKMQ$	1	0	0	0	0	0	0	0	1	1	1	0	1	0	0	0	1
$B = JKLNQ$	0	1	0	0	0	0	0	0	0	1	1	1	0	1	0	0	1
$C = KLMOQ$	0	0	1	0	0	0	0	0	0	0	1	1	1	0	1	0	1
$D = LMNPQ$	0	0	0	1	0	0	0	0	0	0	0	1	1	1	0	1	1
$E = IMNOQ$	0	0	0	0	1	0	0	0	1	0	0	0	1	1	1	0	1
$F = JNOPQ$	0	0	0	0	0	1	0	0	0	1	0	0	0	1	1	1	1
$G = IKOPQ$	0	0	0	0	0	0	1	0	1	0	1	0	0	0	1	1	1
$H = IJLPQ$	0	0	0	0	0	0	0	1	1	1	0	1	0	0	0	1	1

TAB. 32 – *Fractions 2^{n-p} d'aberration minimum ou presque, $p = 8$, $n - p \leq 9$*

rel.déf. — $n = 15, \quad p = 9$

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
$A = JKLMNO$	1	0	0	0	0	0	0	0	0	1	1	1	1	1	1
$B = JKLM$	0	1	0	0	0	0	0	0	0	1	1	1	1	0	0
$C = JKLN$	0	0	1	0	0	0	0	0	0	1	1	1	0	1	0
$D = JMO$	0	0	0	1	0	0	0	0	0	1	0	0	1	0	1
$E = JNO$	0	0	0	0	1	0	0	0	0	1	0	0	0	1	1
$F = KMO$	0	0	0	0	0	1	0	0	0	0	1	0	1	0	1
$G = KNO$	0	0	0	0	0	0	1	0	0	0	1	0	0	1	1
$H = LMO$	0	0	0	0	0	0	0	1	0	0	0	1	1	0	1
$I = LNO$	0	0	0	0	0	0	0	0	1	0	0	1	0	1	1

rel.déf. — $n = 16, \quad p = 9$

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
$A = JKLN$	1	0	0	0	0	0	0	0	0	1	1	1	0	1	0	0
$B = KLMO$	0	1	0	0	0	0	0	0	0	0	1	1	1	0	1	0
$C = LMNP$	0	0	1	0	0	0	0	0	0	0	0	1	1	1	0	1
$D = MNO$	0	0	0	1	0	0	0	0	0	0	0	0	1	1	1	0
$E = JNOP$	0	0	0	0	1	0	0	0	0	1	0	0	0	1	1	1
$F = KOP$	0	0	0	0	0	1	0	0	0	0	1	0	0	0	1	1
$G = JLP$	0	0	0	0	0	0	1	0	0	1	0	1	0	0	0	1
$H = JKM$	0	0	0	0	0	0	0	1	0	1	1	0	1	0	0	0
$I = JKLMNOP$	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1

rel.déf. — $n = 17, \quad p = 9$

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
$A = JKLN$	1	0	0	0	0	0	0	0	0	1	1	1	0	1	0	0	0
$B = KLMO$	0	1	0	0	0	0	0	0	0	0	1	1	1	0	1	0	0
$C = LMNP$	0	0	1	0	0	0	0	0	0	0	0	1	1	1	0	1	0
$D = MNOQ$	0	0	0	1	0	0	0	0	0	0	0	0	1	1	1	0	1
$E = JNOP$	0	0	0	0	1	0	0	0	0	1	0	0	0	1	1	1	0
$F = KOPQ$	0	0	0	0	0	1	0	0	0	0	1	0	0	0	1	1	1
$G = JLPQ$	0	0	0	0	0	0	1	0	0	1	0	1	0	0	0	1	1
$H = JKMQ$	0	0	0	0	0	0	0	1	0	1	1	0	1	0	0	0	1
$I = JKLMNOPQ$	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1

rel.déf. — $n = 18, \quad p = 9$

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
$A = JKLNR$	1	0	0	0	0	0	0	0	0	1	1	1	0	1	0	0	0	1
$B = KLMOR$	0	1	0	0	0	0	0	0	0	0	1	1	1	0	1	0	0	1
$C = LMNPR$	0	0	1	0	0	0	0	0	0	0	0	1	1	1	0	1	0	1
$D = MNOQR$	0	0	0	1	0	0	0	0	0	0	0	0	1	1	1	0	1	1
$E = JNOPR$	0	0	0	0	1	0	0	0	0	1	0	0	0	1	1	1	0	1
$F = KOPQR$	0	0	0	0	0	1	0	0	0	0	1	0	0	0	1	1	1	1
$G = JLPQR$	0	0	0	0	0	0	1	0	0	1	0	1	0	0	0	1	1	1
$H = JKMQR$	0	0	0	0	0	0	0	1	0	1	1	0	1	0	0	0	1	1
$I = JKLMNOPQR$	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1

TAB. 33 – *Fractions 2^{n-p} d'aberration minimum ou presque, $p = 9$, $n - p \leq 9$*

1.12.11 Autres méthodes de choix

Un algorithme de Draper et Mitchell (1968), décrit de façon détaillée dans Kobilinsky (1997), permet d'obtenir tous les types de fraction 2^{n-p} de résolution donnée, à une permutation des facteurs près. Cette recherche conduit par exemple au 5 types de fraction 2^{9-4} de résolution 4 du tableau 34. Les propriétés de deux de ces fractions, obtenues par

	Définition				Profil
1	$F = ABCD$	$G = CDE$	$H = BDE$	$I = ADE$	$4_6 5_8 8_1$
2	$F = ABCD$	$G = CDE$	$H = BDE$	$I = BCE$	$4_7 5_7 9_1$
3	$F = ABCDE$	$G = ABC$	$H = BCD$	$I = ABE$	$4_9 6_6$
4	$F = ABCDE$	$G = ABC$	$H = BCD$	$I = BCE$	$4_{10} 6_4 8_1$
5	$F = ABC$	$G = BCD$	$H = ABD$	$I = ACD$	$4_{14} 8_1$

TAB. 34 – *Fractions 2^{9-4} régulières de résolution 4*

Profil : une notation telle que $4_6 5_8 8_1$ indique 6 mots de longueur 4, 8 de longueur 5 et 1 de longueur 8

PLANOR en introduisant pour chacune sa définition dans la case des *facteurs prédéfinis*, sont reportées dans le tableau 35. L'examen attentif de ces différentes fractions et de leur propriétés peut permettre de sélectionner une fraction appropriée et de choisir de façon judicieuse la correspondance entre les facteurs réels et les lettres, comme cela est illustré dans l'exercice suivant.

Exercice 15 *Retrouver avec PLANOR les propriétés des fractions du tableau 34. Quelles formes ont les interactions estimables dans les fractions* 1, 2, 5. *Quelle similarité voyez vous entre les fractions* 1 *et* 5 *? Pourquoi la* 1 *est-elle préférable à la* 5. *Si on utilise la*

	fraction 1	fraction 2
facteurs définis	$F = ABCD$, $G = CDE$, $H = BDE$, $I = ADE$	$F = ABCD$, $G = CDE$, $H = BDE$, $I = BCE$
mots de définition	$\mathbf{1}$, $EFGHI$, $ACGI$, $ACEFH$, $CDEG$, $CDFHI$, $ADEI$, $ADFGH$, $BCEFI$, $BCGH$, $ABEFG$, $ABHI$, $BDFGI$, $BDEH$, $ABCDF$, $ABCDEGHI$	$\mathbf{1}$, $EGHI$, $BDGI$, $BDEH$, $CDHI$, $CDEG$, $BCGH$, $BCEI$, $ADFGH$, $ADEFI$, $ABFHI$, $ABEFG$, $ACFGI$, $ACEFH$, $ABCDF$, $ABCDEFGHI$
profil	$4_6 5_8 8_1$	$4_7 5_7 9_1$
effets non confondus	A, B, C, D, E, F, G, H, I, AF, BF, CF, DF, EF, FG, FH, FI	A, B, C, D, E, F, G, H, I, AB, AC, AD, AE, AF, AG, AH, AI, BF, CF, DF, EF, FG, FH, FI
effets confondus	(DE,CG,AI,BH),(AB,HI),(AC,GI), (BC,GH),(AD,EI),(BD,EH),(CD,EG), (AE,DI),(BE,DH),(CE,DG),(BI,AH), (AG,CI),(BG,CH)	(BC,GH,EI),(BD,GI,EH),(CD,HI,EG), (BE,DH,CI),(CE,DG,BI),(DE,BH,CG), (CH,BG,DI)
ddl erreur	1	0

TAB. 35 – *Etude fractions régulières* 1, 2 *du tab. 34*

fraction 2*, qui permet d'estimer un nombre maximum d'interactions, à quelles lettres est-il souhaitable d'associer les facteurs les plus importants?*

La notion d'aberration minimum ne privilégie a priori aucun facteur. Dans certains cas, on souhaite d'emblée privilégier certains facteurs et les interactions dans lesquelles ils apparaissent. Une recherche spécifique à l'aide d'un programme tel que PLANOR peut alors permettre de trouvé une fraction appropriée. L'exemple considéré dans l'exercice suivant permet d'illustrer cela plus précisément.

Exercice 16 *On veut étudier l'influence sur la cohésion du fromage de Comté frais de* 11 *facteurs relatifs au procédé de fabrication, étudiés chacun à* 2 *niveaux. On peut fabriquer pour cela* 64 *microcomtés. Donner les relations de définition d'une fraction d'aberration minimum? Quel est son profil? Précisez les groupes d'interactions de deux facteurs confondues entre elles. Combien y-a-t il d'interactions de deux facteurs confondues et non confondues? Peut-on trouver une fraction de résolution* 4 *estimant d'avantage d'interactions de deux facteurs? Y a-t-il un facteur privilégié en ce sens que chacune de ses interactions avec un autre facteur est estimable?*

Les mots de définition de 4 lettres de la fraction d'aberration minimum donnée par le tableau 28 sont, en affectant les lettres A, B, C, D, E, F aux facteurs de base et G, H, I, J, K aux facteurs définis:

$$DEJK, \quad CEIK, \quad CDIJ, \quad ABGH\ .$$

Ces mots contiennent 10 des 11 facteurs. Chacun de ces 10 facteurs apparait donc au minimum dans trois interactions confondues et son effet principal se confond avec au minimum une interaction de trois facteurs.

Si on pense qu'il y a plusieurs facteurs très importants, que l'on veut pouvoir estimer toutes les interactions où ils apparaissent et disposer d'une estimation aussi robuste que possible de leurs effets principaux, on peut préférer un plan ayant peut-être davantage de mots de 4 lettres mais où le nombre de lettres apparaissant dans ces mots est plus faible. Une recherche automatique telle que celle effectuée par le programme PLANOR montre qu'on peut réduire à 8 le nombre de lettres figurant dans les mots de 4 lettres et permet d'obtenir les deux types de fraction du tableau 36 qui rendent estimables toutes les interactions entre l'un des 3 facteurs A, B, C et chacun des 10 autres facteurs. Les deux fractions confondent toutes deux 28 interactions de deux facteurs. Cependant la seconde apparaît préférable parce que les groupes d'interactions confondues y sont de plus petite taille, ce qui peut faciliter beaucoup l'interprétation si les f.e.b correspondantes sont significatives. En contrepartie, le nombre de degrés de liberté résiduels, correspondant aux f.e.b qui n'incluent que des interactions de 3 facteurs ou plus, est inférieur dans la fraction 2. Il y a 12 degrés de liberté pour estimer la variance résiduelle dans la fraction 2 contre 18 dans la fraction 1. Mais cette différence n'apparaît pas fondamentale, surtout si l'on utilise pour l'interprétation les techniques décrites dans Kobilinsky (1997), §3.3.

	Type 1	Type 2
relations de définition	$G = ABCDEF$ $H = ABCE$ $I = ABCF$ $J = DEF$ $K = ABCD$	$G = CDEF$ $H = ABCF$ $I = ABDEF$ $J = ABCE$ $K = ABCD$
Mots de longueur 4	$DHIJ; EGHJ; DEGI; EFHI; DEFJ;$ $FGIJ; DFGH; EIJK; DEHK; GHIK;$ $DGJK; FHJK; DFIK; EFGK$	$EGIJ; EFHJ; FGHI;$ $DEJK; DGIK; DFHK$
Longueur des mots	$W_4 = 14, W_5 = 4, W_7 = 8$ $W_8 = 1, W_9 = 4$ $(4_{14}5_4 7_8 8_1 9_4)$	$W_4 = 6, W_5 = 12, W_6 = 8$ $W_8 = 1, W_9 = 4$ $(4_6 5_{12} 6_8 8_1 9_4)$
Interactions confondues	$(EH,GJ,FI,DK),(DE,GI,FJ,HK),$ $(DH,IJ,FG,EK),(DF,EJ,GH,IK),$ $(DI,HJ,EG,FK),(EF,HI,DJ,GK),$ (FH,EI,DG,JK)	$(EJ,GI,FH,DK),(DE,JK),$ $(GH,FI),(HI,FG),(DJ,EK),$ $(DF,HK),(EI,GJ),(IJ,EG),$ $(DH,FK),(EF,HJ),(GK,DI),$ $(IK,DG),(FJ,EH)$
Effets non confondus	$A; B; C; D; E; F; G; H; I; J; K; AB; AC; AD; AE; AF; AG; AH; AI; AJ; AK$ $BC; BD; BE; BF; BG; BH; BI; BJ; BK; CD; CE; CF; CG; CH; CI; CJ; CK$	
d.d.l résiduels	18	12

TAB. 36 – *Fractions* 2^{11-5} *de résolution* 4 *privilégiant les* 3 *facteurs* A, B, C

Remarques.

- Le mode de recherche utilisé par PLANOR ne garantit pas que les deux types de fraction obtenus sont les seuls possibles.
- Les trois facteurs A, B, C préservés figurent parmi les facteurs de base. La recherche échoue si l'on demande que soit préservés trois des facteurs définis (G, H, I par exemple)

2 Exercices divers

(Provenance : cours ISAB de Christine Durier)

Exercice 17 *Trouver un plan d'expérience pour* 6 *facteurs comportant un nombre d'expériences* *minimum et permettant d'estimer tous les effets principaux et interactions de* 2 *de facteurs.*

Exercice 18 *Soit le plan factoriel* 2^{5-2} *défini par :*$ABCD = 1$, $ACE = 1$. *On postule que*

- *les facteurs* A, B, C *n'interagissent pas entre eux,*
- *les facteurs* C, D, E *n'interagissent pas entre eux*
- *les interactions de* 3 *facteurs et plus sont négligeables*

Quels sont les effets principaux et interactions de 2 *facteurs estimables ?*

Exercice 19 *Pour la construction d'un quart de plan* 2^5, *deux définitions sont proposées :*

$$\begin{array}{lll} a. & ABCDE = 1, & BCDE = 1 \\ b. & ABC = 1, & BCDE = 1 \end{array}$$

Préférez-vous le plan construit avec a. ou b. ? Pourquoi ?

Exercice 20 *On étudie* 6 *facteurs* A, B, C, D, E, F.

Un plan 2^{6-3} *est défini par :* $1 = ABD = ACE = BCF$.

Il est complété par un plan identique mais dans lequel les -1 *et* $+1$ *sont inversés pour les facteurs* A *et* B. *En supposant négligeables les interactions de* 3 *facteurs et plus, quels sont les effets confondus dans le plan à* 16 *traitements ainsi défini ?*

Exercice 21 *Un expérimentateur commence une étude par un plan* 2^{5-2} *défini par* $D = ABC$ *et* $E = BC$. *Plus tard, il complète ce plan avec le plan identique dans lequel tous les signes sont inversés.*

a. Quelle est la résolution du plan global ainsi défini ? Comment est-il défini ?

b. Pouvez-vous suggérer un plan meilleur, à 16 *traitements, si l'expérimentateur avait pu savoir à l'avance que 16 essais seraient nécessaires ?*

(Provenance : cours ENSAE de H. Monod, Christine Durier)

Exercice 22 *On veut étudier 5 facteurs à 2 niveaux notés* A, B, C, D, E *avec une première expérience en seulement* 8 *essais.*

1. *Quelle est la meilleure résolution que l'on peut obtenir ?*

On veut estimer les effets principaux de tous les facteurs, en supposant que toutes les interactions sauf $B.D$, $B.E$, $C.D$, $C.E$ *sont nulles.*

2. *Préciser (sans nécessairement les énumérer) les effets factoriels qui ne doivent pas être confondus avec la moyenne. Quelles sont les interactions entre* 3 *facteurs qui peuvent être confondues avec la moyenne ?*
3. *En déduire une solution pour la fraction en précisant ses contrastes de définition.*

REFERENCES

Chen J. (1992). Some results on 2^{n-k} fractional factorial designs and search for minimum aberration designs. *Ann. Statist.*,**20**, 2124-2141.

Chen J. and Wu C.F.J. (1991). Some results on s^{n-k} fractional factorial designs with minimum aberration or optimal moments. *Ann. Statist.*,**19**, 1028-1041.

Draper N.R. and Mitchell T.J. (1968). Construction of the set of 256-run designs of resolution ≥ 5 and the set of even 512-run designs of resolution ≥ 6 with special reference to the unique saturated designs. *Ann. Math. Statist.*, **39**, 246-255.

Franklin M.F. (1984). Constructing Tables of Minimum Aberration p^{n-m} Designs. *Technometrics*, **26**, 225-232.

John P.W.M (1971). Statistical Designs and analysis of experiments. Macmillan, New-York.

Kobilinsky A. (1990). Complex linear models and cyclic designs. *Linear Algebra and Applications*, **127**, 227-282.

Robillard P. (1968). Combinatorial problems in the theory of factorial designs and error correcting codes. Inst. Statistics Mimeo Ser. 594. Univ. North Carolina, Chapel Hill.

Statistical Engineering Laboratory (1957, 1962). Fractional Factorial Experiment Designs for factors at two levels. *National Bureau of Standards. Applied Mathematics Series*, . (Reproduit dans McLean and Anderson [1984]).

McLean R.A. and Anderson V.L. (1984). Applied factorial and fractional designs. Marcel Dekker, New-York, 373p.

Table des matières

Table des figures

Liste des tableaux