



HAL
open science

CHAPITRE XI: Outils de Simulation des modes opératoires & Plans d'expériences

Mohammed Tamali

► **To cite this version:**

Mohammed Tamali. CHAPITRE XI: Outils de Simulation des modes opératoires & Plans d'expériences: Concepts de base & fondements. Master. Algérie. 2013. cel-01445451

HAL Id: cel-01445451

<https://hal.science/cel-01445451>

Submitted on 24 Jan 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License

Université de Béchar
Laboratoire des Études Énergétiques en Zones Arides
Équipe Modélisation & Simulation des Systèmes

CHAPITRE XI :
Outils de Simulation des
modes opératoires
Plans d'expériences
Concepts de base & fondements.

Experimental Design Process



Version
2.3.0

Cours réalisé par : Prof. TAMALI Mohammed,

<http://www.univ-bechar.dz/mtamali>

Université de Béchar | FS&T
(ENERGARID Lab./SimulIA Team)



pyDOE
Design of Experiments for Python





Presentation

The University of Bechar was born in 1986 as the National Institutes of Higher Education (INES) in 1992 it becomes University Center and on January 07, 2007, it was officially declared as a University. Since then, many Research Teams have seen the day. In 2011, The Laboratory for Energy Systems Studies Applied to Arid Zones was run by a group of young and well motivated researchers (7 Research teams) to solve real problems affecting arid zones, Simulia Team is one of them in the same laboratory. The Workload of Simulia concern studies and applications of modeling and simulation of systems in Arid Areas.

Research areas:

- Energy & Environment (Modeling & Simulation)
- Application of heat in arid zones
- Energy economy.
- Mapping and development of resources in arid zones.
- SIMULIA for the task in the short term, to develop the computer code for modeling and simulation which can be accessed online.

Website of the laboratory team: <http://energarid.wordpress.com/>



Généralités & Présentations

Concepts fondamentaux & Définitions

Rappel des notions de la statistique

Facteur & Modalités

Outils Python pour le traitement des PDE (DOE)

Modélisation & mise en équations

Plans factoriels complet à deux niveaux 2^k

Plans factoriels fractionnaires à deux niveaux 2^{k-p}

Erreurs expérimentales

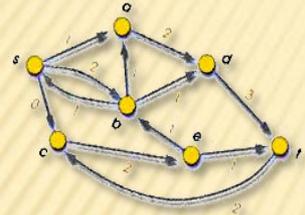
Plan de second degré

Analyse de la variance

Plan de mélange

Conclusions & Références bibliographiques

Généralités & Présentations



Paramètres multimodaux

Les méthodologies utilisées par les humains, en rapport, avec les tentatives de compréhension des phénomènes physiques qui nous entourent, donnent un contrecoup général de la complexité de ces mêmes systèmes que nous manipulons et prenons comme sujets dans nos études.

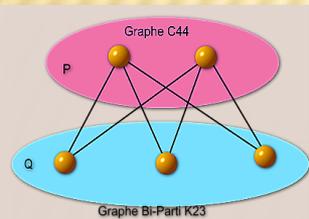
Le niveau de complexité des systèmes **est élevé**, à un niveau où toutes les tentatives ou essais de lancement de procédés expérimentaux laissent et obligent à considérer des erreurs. Encore plus, les effets tangents. Selon la théorie de l'évaluation des performances, l'exigence 'comprendre' le système n'a de réponse que si :

- Nous avons tellement **d'informations** que les recommandations des études ultérieures seront satisfaites,
- Nous avons des **références**, avec quoi peut-on comparer,
- Nous avons un **historique** susceptible d'être retracé,
- Il y a une possibilité pour faire de **l'expérimentation**.

Les trois premiers cas satisfont à eux-mêmes. Si telle est la situation, ils nous clarifient l'image. Le quatrième critère exige que l'expérience se fasse effectivement pour que toutes les questions, relatives à un problème donné, soient élucidées. Le **domaine de définition du modèle** régissant le système étudié est plus ou moins profond que ses variables se meuvent d'une manière continue ou discrète dans les espaces position/temps.

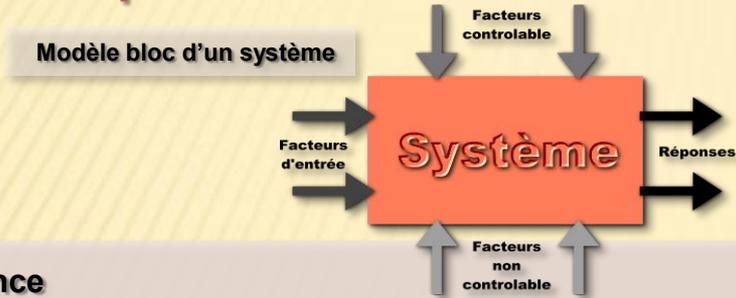
Ces variables sont les **facteurs** du systèmes et peuvent évoluer selon des **modalités** changeantes.

Graphe Biparti



Graphe Bi-Parti K23

Concepts fondamentaux & Définitions



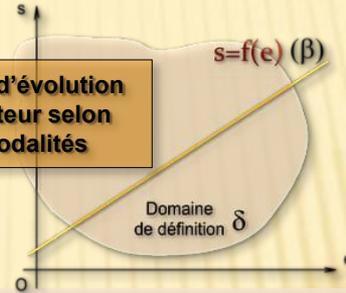
Expérience

n. f. Expérimenter, **acte** de procéder à des **essais effectifs** dans un **lieu** destiné à cette tâche (laboratoire) et pour un but purement **scientifique**.

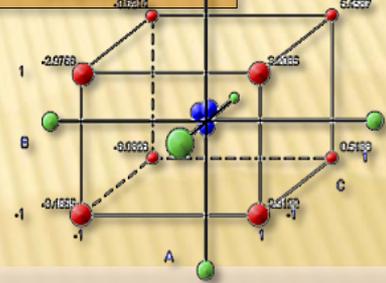
Selon la complexité de la composition du système, on est appelé à faire **beaucoup de tests** (m expériences) pour un même **scénario** et ceci, selon la dépendance (ou non) entre les **facteurs** régissant le système, sujet de l'expérience.

Cette procédure est d'autant plus combinatoire que le nombre de cas à vérifier est beaucoup plus grand ($m \Rightarrow \infty$). Les **plans d'expériences** permettent d'organiser au mieux les essais qui accompagnent une recherche scientifique ou des études industrielles [1]. Ils sont applicables à de nombreuses disciplines et à toutes les industries à partir du moment où l'on recherche le lien qui existe entre une grandeur d'intérêt, y et des variables, x_i . Il faut penser aux plans d'expériences si l'on s'intéresse à une fonction du type $y = f(x_i)$. Avec les plans d'expériences on obtient le maximum de renseignements avec le minimum d'expériences.

Courbes d'évolution d'un facteur selon des modalités



Courbes d'évolution d'un phénomène physique



Modélisation mathématique :

Procédures mathématique permettant de choisir un équivalent (Modèle) à un système donné.

Facteur, Modalité & Plan

Facteur : les composantes (facteurs) d'un système sont les éléments dont il dépend. Ce sont ses variables.

Modalité (Niveau) : Niveau appréciable que peut prendre un facteur relativement à une situation précise du système.

Plan : Une composition expérimentale visant à faciliter la tâche de l'expérimentateur en lui présentant une méthodologie finie pour entreprendre ses études et essais.

Espace expérimental : Le domaine de variabilité des réponse à toutes les valeurs du facteur étudié selon des modalités.

Concepts fondamentaux & Définitions

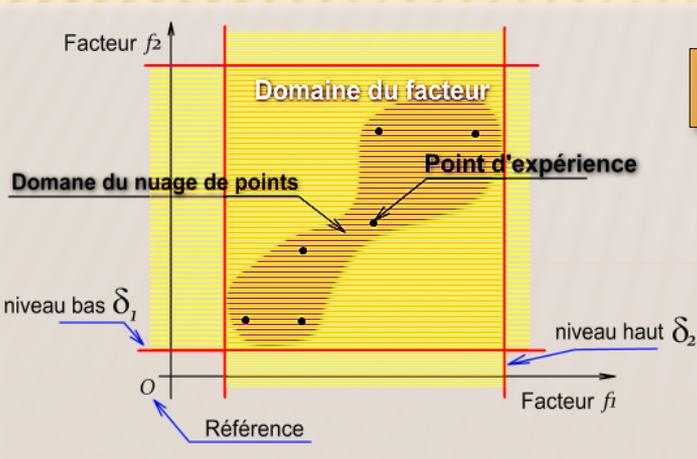
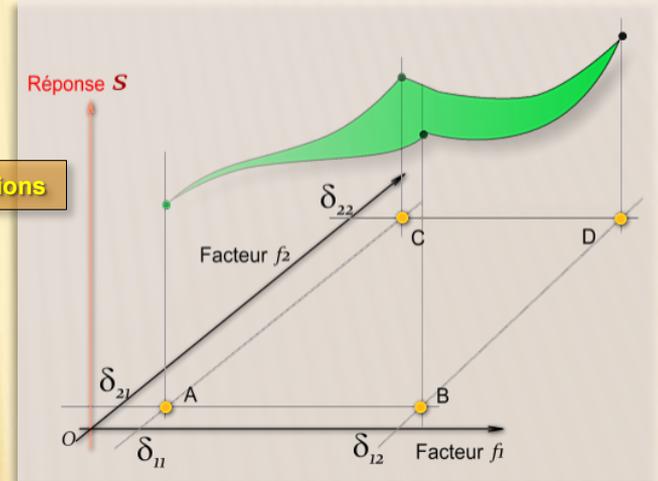


Situation d'un seul facteur

Réponse aux interactions

L'expérience peut être menée selon beaucoup de manière, le phénomène étudié est au centre des préoccupations. Le cas le plus simple est celui où la variation d'un seul paramètre est considéré.

Le système dépend, fondamentalement, de ses composantes intrinsèques qui varient dans le temps et l'espace physiques. Le cas général à considérer est; quand plus d'une composante rentre en **interaction**.

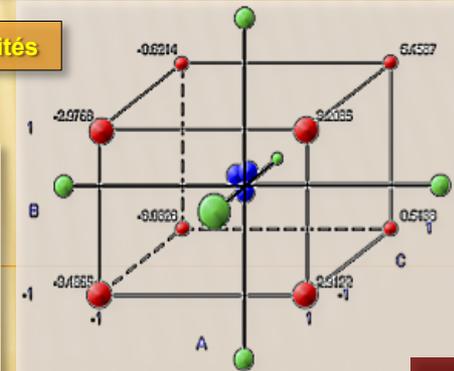


Domaine de variabilité de deux facteurs

Facteurs & Modalités

La réponse du système étudié dépend essentiellement du type de considérations des rattachements à remarquer dans les interactions afin qu'une réponse du système ne soit ressentie.

L'observation des interactions **doubles** est la plus simple à exécuter. Les systèmes physiques sont, par défaut, **trop complexes**.



Concepts fondamentaux & Définitions

La réponse du système selon les modes de variations des facteurs, sujet de l'expérimentation, est assimilable à une fonction y pouvant être déterminée par un développement limité de Taylor-Mac Laurin.

$$y = a_0 + \sum a_i \cdot x_i + \sum \sum a_{ij} \cdot x_i \cdot x_j + \sum \sum a_{ij} \cdot x_i^2 + \dots + \sum \sum \dots \sum a_{ijklm\dots z} \cdot x_i \cdot x_j \cdot x_k \dots x_z$$

Equation du **modèle à priori**.

Les dérivées sont constantes. y est mesurée (grandeur d'intérêt) pendant l'expérimentation, selon un scénario choisi.

x_i représente le niveau attribué au **facteur i** par l'expérimentateur pour exécuter un scénario (vue expérimentale). On supposera que cette valeur est bien connue.

Les coefficients a_i, a_{ij}, \dots , sont inconnus et représentent d'ailleurs **les paramètres à déterminer**, les scénarios sont bien choisis. Les résultats de l'expérience permettent de connaître ces valeurs.

L'intérêt majeur de cette manipulation (modéliser la réponse y par un polynôme) c'est de pouvoir extrapoler et calculer toutes les réponses du domaine d'étude sans être obligé de faire des expériences en plus. En d'autres termes, **optimiser le PLAN D'EXPERIENCES**.

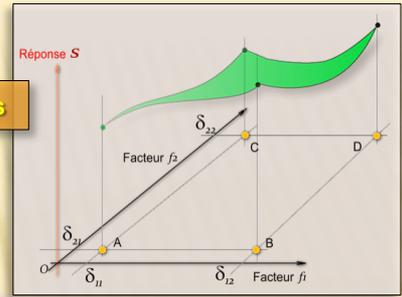
Les ajustements probables dans les valeurs des facteurs choisis et l'aspect aléatoire de la valeur de la réponse, susceptible d'être obtenue après des essais laisse à remarquer des erreurs (**erreurs expérimentales**) non considérées dans le modèle à priori. Le modèle est réécrit de façon suivante :

$$y = a_0 + \sum a_i \cdot x_i + \sum \sum a_{ij} \cdot x_i \cdot x_j + \sum \sum a_{ij} \cdot x_i^2 + \dots + \sum \sum \dots \sum a_{ijklm\dots z} \cdot x_i \cdot x_j \cdot x_k \dots x_z + e$$

Equation du **modèle de l'expérimentateur (e : erreur)**.

Ce modèle suppose la réponse du système sensible à toute variation des composantes

Facteurs & Modalités



La valeur d'un point (concours des estimations de facteurs x_i) permet d'obtenir une valeur de la réponse y . Cette dernière est modélisée par un polynôme dont les coefficients ne sont connus, donc à déterminer. Le scénario du plan d'expériences nous mène à un système de m équations (d'où m essais) à p inconnues (d'où p coefficients dans le modèle choisi à priori). Ce système s'écrit alors : $Y = A \cdot X + E$.

Y Vecteur des m composantes de la réponse.

X Matrice de calcul, ou matrice du modèle, qui dépend des points expérimentaux choisis.

A Vecteur des coefficients dans le modèle.

E Vecteur des écarts et erreurs commises lors de l'expérience.

Le système qui en découle possède (m) équations et ($p + n$) inconnues (plus d'inconnues que d'équations).

Concepts fondamentaux & Définitions

Les coefficients sont obtenus, selon une méthode de régression basée sur le critère des moindres carrés :

$$\hat{A} = (X^T \cdot X)^{-1} \cdot X^T \cdot Y$$

X^T étant la matrice transposée de X , $X^T \cdot X$ est dite matrice d'information et $(X^T \cdot X)^{-1}$ matrice de dispersion. \hat{A} Vecteur des Estimations des coefficients du modèle de l'expérimentateur.

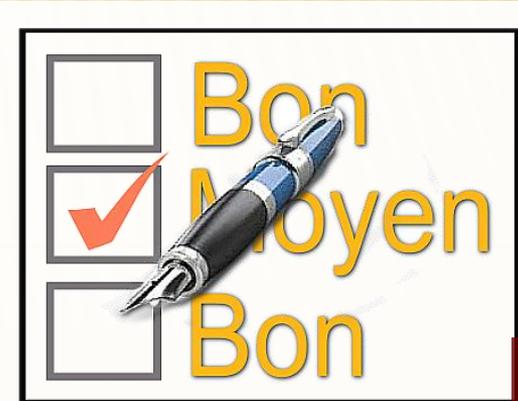
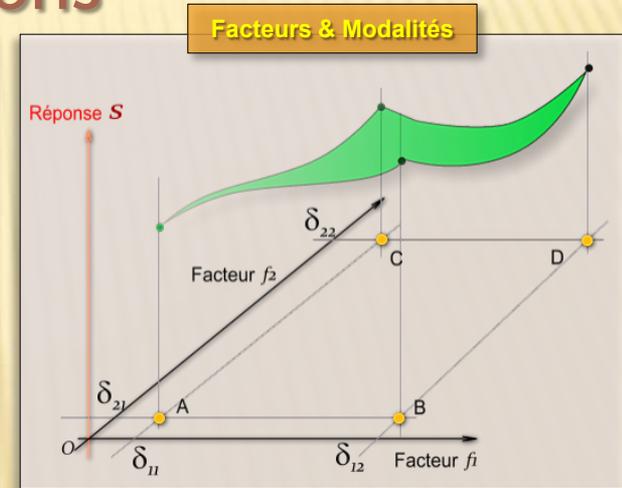
L'utilisation d'un plan d'expérience donne alors une stratégie dans le choix des méthodes d'expérimentation. Le succès des plans d'expériences dans la recherche et l'industrie est lié au besoin de compétitivité des entreprises : ils permettent une amélioration de la qualité et une réduction des coûts.



Ronald Aylmer Fisher
1890-1962

La méthode des plans d'expériences a été mise au point au début du XX^e siècle, dans les années 1920, par Ronald A. Fisher (Biologiste & Statisticien britannique), dans le cadre d'études agronomiques. Elle a pris un essor considérable avec le développement de l'informatique et la puissance de calcul qui l'accompagne. La grande nouveauté de la méthode des plans d'expériences est qu'elle propose une expérimentation factorielle, c'est-à-dire que tous les facteurs varient simultanément. Le traitement des résultats se fait à l'aide de la régression linéaire multiple et l'analyse de variance.

D'une bonne évaluation découle une bonne estimation et par suite une confiance.



Rappel des notions de la statistique

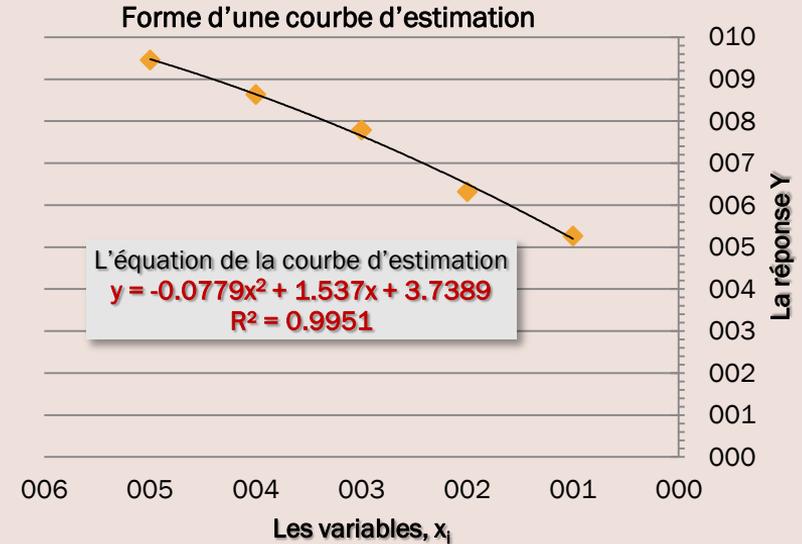
Régression : méthode utilisée en estimation de la forme la plus proche à l'évolution représentée par un nuage de points issus de l'expérience. A partir d'un ensemble de valeurs expérimentales, que nous devons représentées par des points sur un graphique, on procède à la détermination de la courbe qui reproduit, d'une manière fidèle, les changements des grandeurs étudiées et de leur réponse. Cette courbe représentera la meilleure estimation des résultats.

La régression linéaire multiple est une méthode d'analyse de données quantitatives. Elle a pour but de mettre en évidence la liaison pouvant exister entre une variable dite **expliquée** notée Y et plusieurs autres variables dites **explicatives** que nous noterons x_1, x_2, \dots, x_k .

On note plusieurs formes de courbes pouvant être utilisées pour équivaloir un nuage de points X à la bonne réponse de l'estimation. La forme de la régression à adopter est celle dont le carré des erreurs, entre l'estimateur Y' (courbe) et les valeurs issues de l'expérience Y , est le plus petit en tout point de X . Elle peut être linéaire ($y=A^T X+B$), Polynomial de degré n ($y=P_n(X)=\sum a_i x^i$), logarithmique ($y=\ln(a.x)+b$) ou autre ...

La forme linéaire est la plus utilisée, vu que le type d'interaction le plus facile à étudier est le plus simple. Les interactions d'ordre supérieur sont difficile à manipuler. On se contente de prendre en considération des interactions multiples mais à un seule variable.

Distribution en nuage de points



Y	x_i
5.265	1.00
6.321	2.00
7.789	3.00
8.635	4.00
9.4569	5.00

Rappel des notions de la statistique

Soit Ω l'univers associé à une expérience aléatoire.

Événement élémentaire : est un événement qui ne contient qu'une seule issue.

Espérance de la loi de probabilité est le nombre μ défini par $\mu = \sum P_i \cdot x_i$, P_i étant la probabilité de l'évènement élémentaire $\{x_i\}$.

Moyenne : pour un ensemble d'expériences faites sur un système, une variable aléatoire est prise/considérée selon un vecteur de N valeurs. La moyenne de ces valeurs est $M_x = (1/N) \cdot \sum x_i$.

Variance : la variance est donnée par $V(X) = \sigma_x^2 = [1/(N-1)] \cdot \sum (x_i - M_x)^2$.

Covariance : pour deux variables aléatoires, la covariance est donnée par la relation; $\sigma_{xy}^2 = [1/(N-1)] \cdot \sum [(y_i - M_y)(x_i - M_x)]$.

Ecart-type : est un réel positif défini par $\sigma = \sqrt{V(X)}$.

Ajustement : c'est toute solution du système des n équations (sans interactions)

$Y_i = a_0 + a_1 x_{i1} + \dots + a_k x_{ik} + e_i$ (avec $i \in [1, 2, \dots, n]$)

Les a_i sont les coefficients du modèle a priori, x_{ik} est la valeur de x_k choisie par l'expérimentateur lors de l'expérience i , Y_i étant la valeur observée. Les e_i sont les résidus (erreur comises lors de l'expérience i).

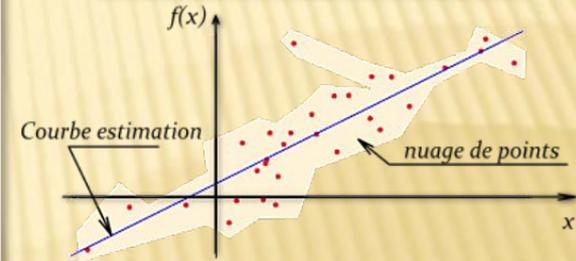
Sous une notation matricielle nous avons $Y_i = A^T \cdot X_i + E_i$, d'où E_i est égal à la différence $(Y_i - A^T \cdot X_i)$.

Un ajustement au sens des moindres carrés laisse voir la relation suivante $\sum e_i^2 = S_{\min}$.

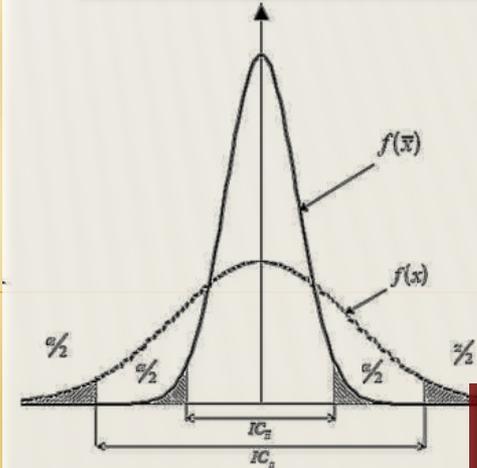
S doit être minimale pour dire que les estimateurs \hat{a}_i sont les bonnes valeurs recherchées de a_i , alors S ne peut être extrémale (min) qu'aux points vérifiant les conditions d'optimalité $\partial S / \partial x_k = 0$.

On a encore pour $\lambda \in \mathbb{R}$; les variances $V(\lambda \cdot X) = \lambda^2 \cdot V(X)$ et $V(X \pm Y) = V(X) + V(Y)$.

Distribution en nuage de points



Distribution en cloche de Gauss



Outils Python pour le traitement des PDE (DOE)

Domaine d'application

Dans l'environnement Python, il existe des librairie spécifique pour le traitement et la manipulation des plan d'expériences.

La librairie désignée est PyDOE

Procédures

Installer PYTHON (2.7-3.4+)

Installer PIP, DSUTILS

Installer NUMPY et SCIPY

Et enfin installer PyDOE



Cette librairie peut être utilisée

En deux manière

- Fenêtre TERMINAL
- Navigateur JUPYTER

A screenshot of a Jupyter Notebook interface. The browser address bar shows 'localhost:8888/notebooks/documents/notebooks/pydoe.ipynb#'. The notebook title is 'pydoe (autosaved)'. The code cells contain:

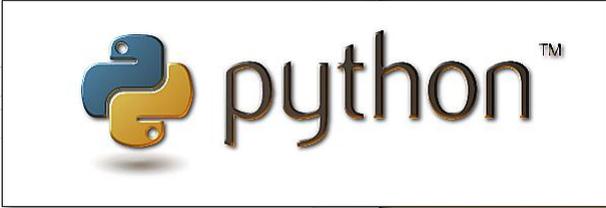
```
In [24]: import pyDOE as doe
import numpy as np

In [31]: levels = [[ 0, 0],
[ 1, 0],
[ 0, 1],
[ 1, 1],
[ 0, 2],
[ 1, 2]]

In [32]: ff2n(3)
Out[32]: array([[ -1., -1., -1.],
[ 1., -1., -1.],
[ -1., 1., -1.],
[ 1., 1., -1.],
[ -1., -1., 1.],
[ 1., -1., 1.],
[ -1., 1., 1.],
[ 1., 1., 1.]])

In [36]: gen = 'a b ab'

In [37]: fracfact('a b ab')
Out[37]: array([[ -1., -1., 1.],
[ 1., -1., -1.],
[ -1., 1., -1.],
[ 1., 1., -1.],
[ -1., -1., 1.],
[ 1., -1., 1.],
[ -1., 1., 1.],
[ 1., 1., 1.]])
```

An orange callout box with the text 'Fenêtre JUPYTER' is overlaid on the right side of the notebook interface.

Modélisation & mise en équations

Domaine d'application

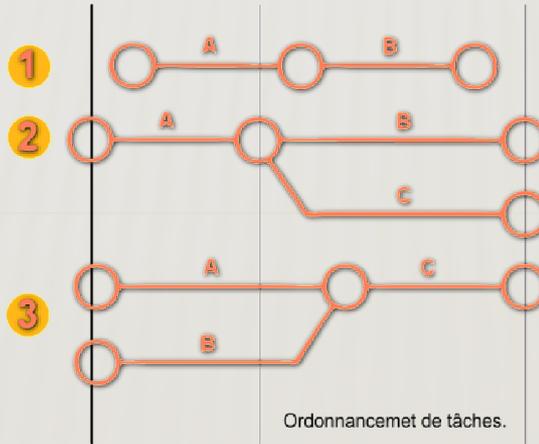
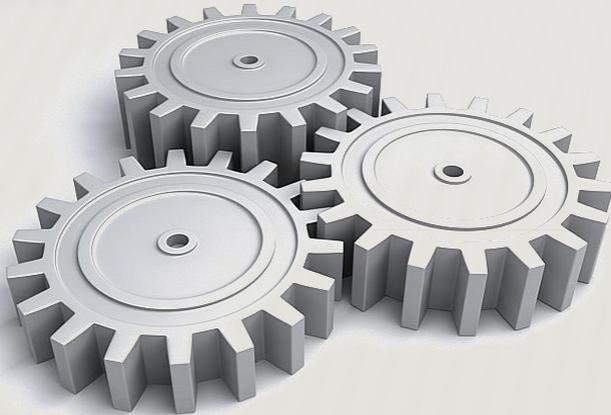
Les plans d'expériences sont utilisés dans les études industrielles en recherche-développement. Ils interviennent dans de nombreux domaines industriels. On peut notamment citer :

- industries chimiques, pétrochimiques et pharmaceutiques
- industries mécaniques et automobiles
- industries métallurgiques

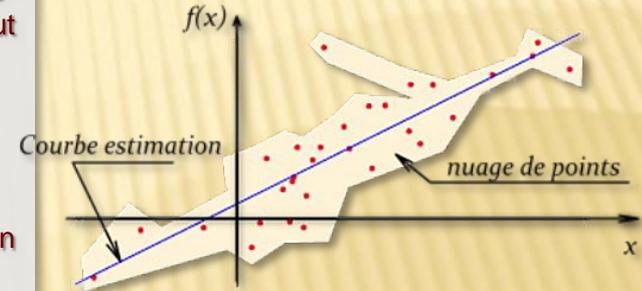
Leur utilisation vise aux buts suivants :

- détermination des facteurs clés dans la conception d'un nouveau produit ou d'un nouveau procédé
- optimisation des réglages d'un procédé de fabrication ou d'un d'appareil de mesure
- prédiction par modélisation du comportement d'un procédé.

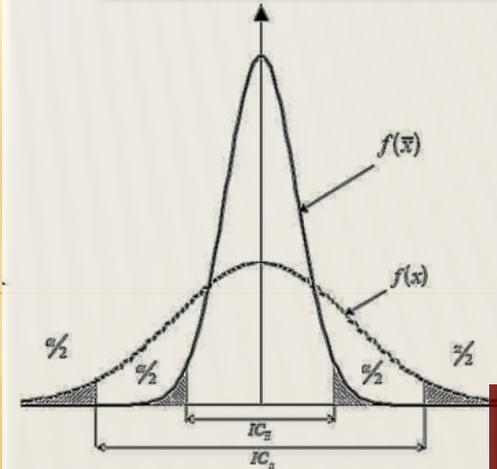
Les plans d'expériences s'inscrivent dans une démarche générale d'amélioration de la qualité.



Distribution en nuage de points



Distribution en cloche de Gauss



Modélisation & mise en équations

Variables codées ou variables centrées réduites :

Nous remplaçons les variables physiques par les variables de notre imagination, dites **codées**, ceci réduira la prise en considération de chaque facteur dans le même domaine de variation que les autres (entre **-1** et **+1**) et de pouvoir ainsi faire la comparaison entre eux et de même pour leurs effets.

Le niveau bas (**-1**) ainsi code la borne (**niveau, modalité**) inférieure du domaine de variation du facteur alors que le niveau haut est codé par (**+1**).

Tableau d'expérimentation et Plan d'expérience

Un tableau d'expérimentation utilise toujours un regroupement des essais à effectuer selon les choix des niveaux des facteurs pris en compte en **valeurs physiques**.

Alors, le plan d'expérience, quant à lui, il reprend le même tableau mais en **valeurs codées** (avec indication de l'échelle de codage).

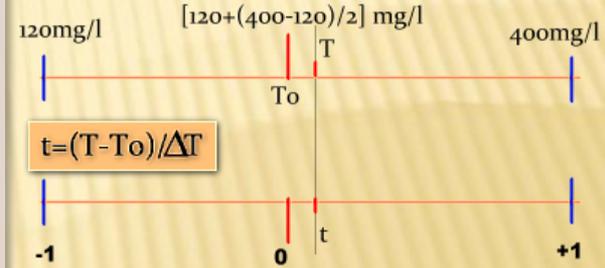
Essai Numéro	Facteur 1 p	Facteur 2 q
01	p_{\min}	q_{\min}
02	p_{\max}	q_{\min}
03	p_{\min}	q_{\max}
04	p_{\max}	q_{\max}

Tableau d'expérimentation
exemple de deux facteurs p et q

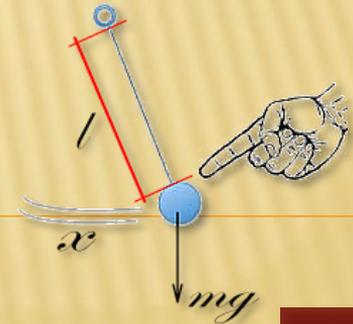
Essai Numéro	Facteur 1 p	Facteur 2 q	Réponse y
01	-1	-1	y_1
02	+1	-1	y_2
03	-1	+1	y_3
04	+1	+1	y_4
Niveau -1	p_{\min}	q_{\min}	
Niveau +1	p_{\max}	q_{\max}	

Plan d'expérience, exemple de deux facteurs p et q.

Variables centrées réduites



$$x'' + \frac{g}{l} \sin x = 0$$



Modélisation & mise en équations

Calcul des coefficients :

Pour un problème à deux facteurs nous aurons comme modèle à priori

$$y = a_0 + a_1 \cdot x_1 + a_2 \cdot x_2 + a_{12} \cdot x_1 \cdot x_2 \text{ (e étant négligée).}$$

Selon le plan d'expérience, comme montré dans le slide précédent, cette même équation peut être écrite selon les quatre cas de figures :

$$y_1 = \hat{a}_0 - \hat{a}_1 - \hat{a}_2 + \hat{a}_{12}$$

$$y_2 = \hat{a}_0 + \hat{a}_1 - \hat{a}_2 - \hat{a}_{12}$$

$$y_3 = \hat{a}_0 - \hat{a}_1 + \hat{a}_2 - \hat{a}_{12}$$

$$y_4 = \hat{a}_0 + \hat{a}_1 + \hat{a}_2 + \hat{a}_{12}$$

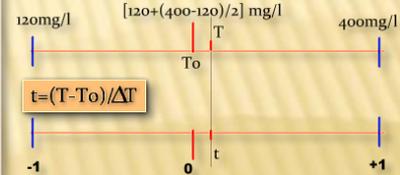
Par un simple calcul manuel, on déterminera les coefficients \hat{a}_i à partir de ces équations. De cette manière, nous aurons la forme de **régression linéaire** qui identifie l'évolution de notre processus selon les variations de ses facteurs étudiés : $y' = \hat{a}_0 + \hat{a}_1 \cdot x_1 + \hat{a}_2 \cdot x_2 + \hat{a}_{12} \cdot x_1 \cdot x_2$

On constate que chaque coefficient \hat{a}_i se calcule en prenant la somme des réponses y_j , chacune d'elles étant affectée du signe de la colonne correspondante à ce coefficient, divisé par le nombre d'essais. La puissance des plans d'expériences réside dans le fait qu'ils permettent la réduction du nombre d'expériences nécessaires pour obtenir une bonne estimation des mécanismes du système étudié. Toutefois, on ne peut se contenter d'effectuer un nombre d'expériences que pour des raisons de diminution des tâches mais un minimum est nécessaire pour atteindre un niveau appréciable et une précision optimale.

Matrice des effets

Le système d'équations à résoudre doit présenter des coefficients devant les inconnues et peut se mettre sous la forme d'une matrice dite **matrice des effets**. Par exemple pour un système avec effet d'interaction ($a_{ij} \neq 0$) de deux facteurs chacun pris à deux modes, nous aurons la matrice (dite d'Hadamard).

Variables centrées réduites



Matrice des effets

$$\begin{bmatrix} +1 & -1 & -1 & +1 \\ +1 & -1 & +1 & -1 \\ +1 & +1 & -1 & -1 \\ +1 & +1 & +1 & +1 \end{bmatrix}$$



Jacques Hadamard
(Mathématicien, Fr)
1865-1963

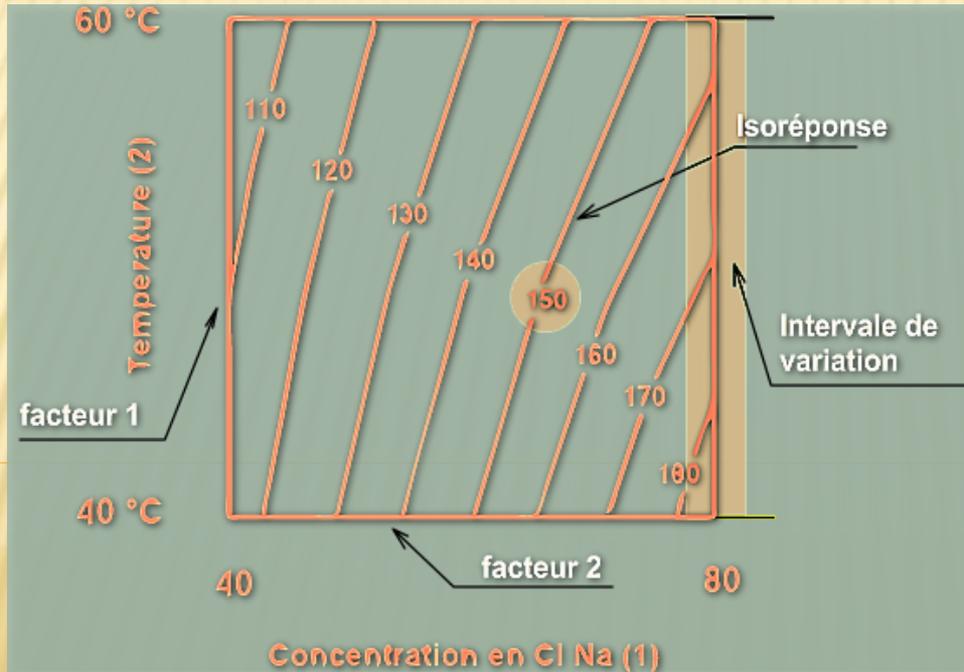
Modélisation & mise en équations

Courbes iso-réponse :

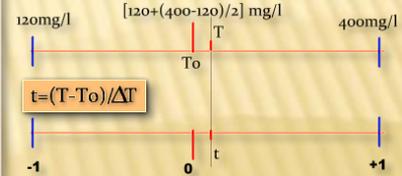
Pour avoir une vue d'ensemble des résultats, on trace les courbes iso-réponse dans le domaine d'étude et ceci pour distinguer comment varie une réponse y selon les variations des variables x_i à l'origine. Elle est facilement traçable pour le cas de nombre de facteurs égale à 2.

On fait varier x_1 dans son domaine en considérons les autre variables constantes (idem pour x_2).

$$x_1 = (1/\hat{a}_1)[y' - (\hat{a}_0 + \hat{a}_2 \cdot x_2 + \hat{a}_{12} \cdot x_1 \cdot x_2)]$$



Variabes centrées réduites



Matrice des effets

+1	-1	-1	+1
+1	-1	+1	-1
+1	+1	-1	-1
+1	+1	+1	+1



Jacques Hadamard
(mathématicien, FR)
1865-1963

Plans factoriels complet à deux niveaux 2^k

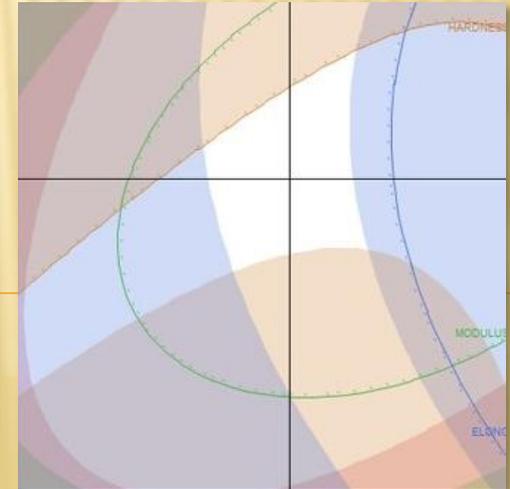
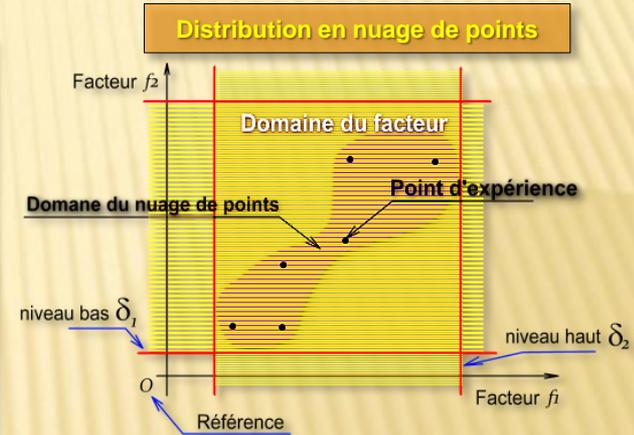
Dans le cadre de ces plans nous nous limiterons à deux niveaux pour chaque facteur considéré à part. Les équations prendrons en considération des formations de niveaux au cours de l'expérimentation. L'application peut se faire pour le cas de tout type de variable (continue ou discrète).

$$y = a_0 + \sum_i a_i x_i + \sum_i \sum_j a_{ij} x_i x_j + e$$

- y est la réponse du système
- x_i représente le niveau attribué au facteur i .
- a_0 est la valeur de la réponse au centre du domaine d'étude.
- a_1 est l'effet (ou effet principal) du facteur 1.
- a_2 est l'effet (ou effet principal) du facteur 2.
- a_{12} est l'interaction entre les facteurs 1 et 2.
- e est l'écart.

La représentation graphique est limitée et elle perd son sens une fois le nombre de variables dépasse 3. De ce fait, une représentation en mode tableau est utilisée sans à ce qu'elle soit dépendante du nombre de variables ni des limites des niveaux à modéliser.

Un **PLAN FACTORIEL COMPLET** est un cas de plan pour lequel toutes les combinaisons possibles d'essais seront réalisées. Le nombre d'expériences N se calcule d'après $N=2^k$ où $k \in \mathbb{N}$ est le nombre de facteurs régissant le système.



Plans factoriels complet à deux niveaux 2^k

2 niveaux/facteur

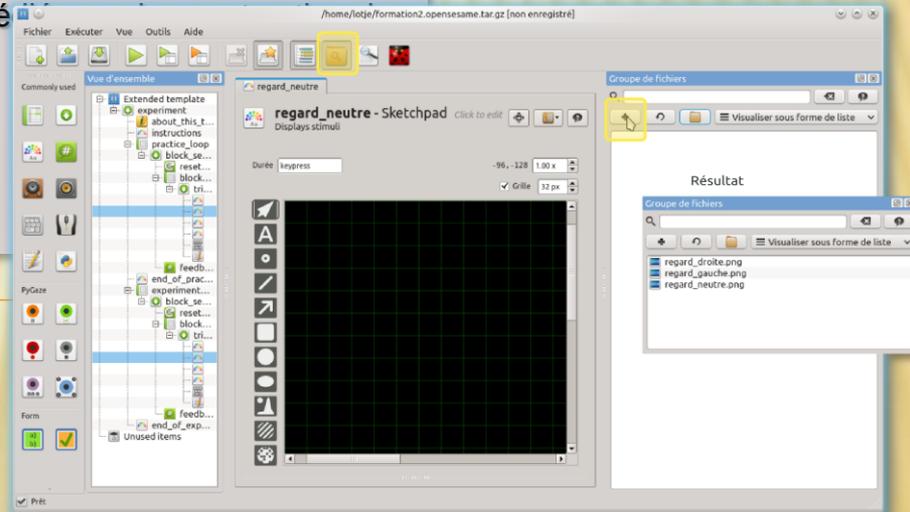
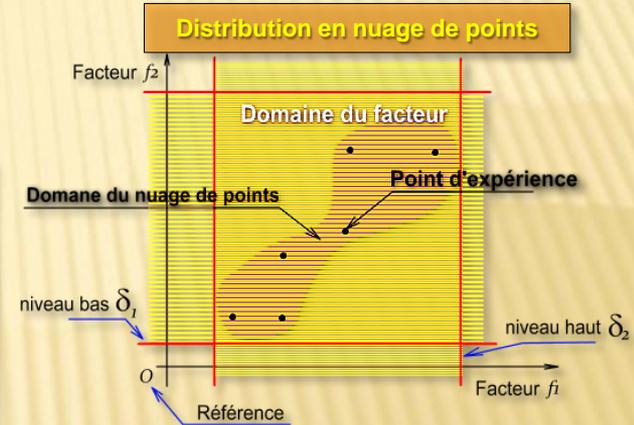
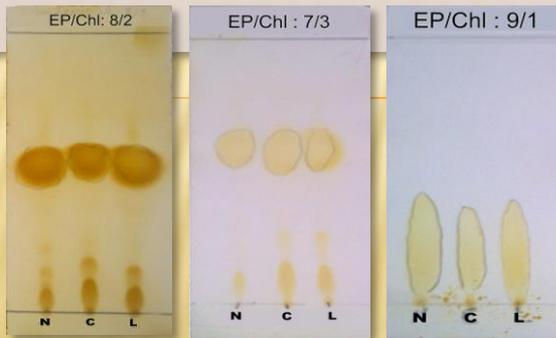
K facteurs

En général, la matrice d'expériences comporte k colonnes pour les facteurs principaux et 2^k lignes soit 2^k essais. Elle se construit selon la règle suivante :

- colonne du 1^{er} facteur: alternance de -1 et +1
- colonne du 2^e facteur: alternance de -1 et +1 de 2 en 2
- colonne du 3^e facteur: alternance de -1 et +1 de 4 en 4
- colonne du 4^e facteur: alternance de -1 et +1 de 8 en 8

et ainsi de suite pour un nombre plus élevé de facteurs.

OpenSesame est un outil graphique, open-source de expériences pour les sciences sociales. Il arbore une interface utilisateur moderne et intuitive qui permet de construire des expériences complexes avec un minimum d'efforts.



Plans factoriels complet à deux niveaux 2^k

Exemple de cas d'étude :

- une étude sur l'usure des pneus (réponse) montre une interaction entre la vitesse et la pression de gonflage (facteurs)
- une étude sur la somnolence a montré une interaction entre la quantité d'alcool ingéré et la quantité d'un médicament particulier absorbé. Le graphique ci-contre montre les valeurs de la somnolence mesurée en unité arbitraire de 0 à 10 en fonction des deux autres facteurs mesurés également avec des unités arbitraires.

On note une forte interaction en remarquant que l'augmentation de somnolence est beaucoup plus forte si la quantité d'alcool ingéré est plus importante. **L'effet du médicament dépend donc de la quantité d'alcool ingéré.**

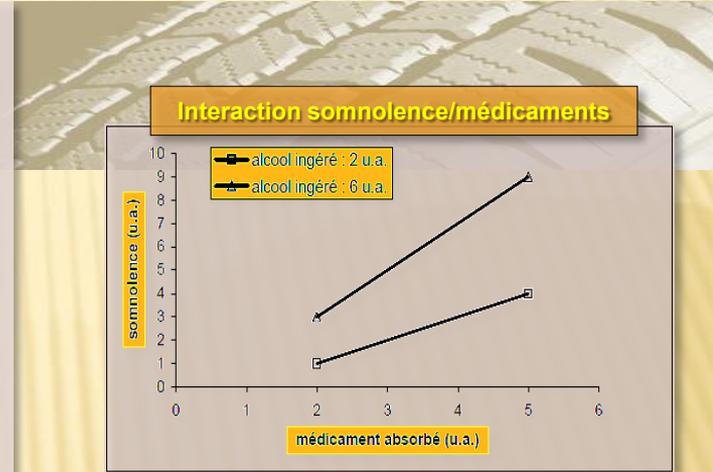
Une application simple est fournie par le plan d'expériences suivant où les calculs peuvent s'effectuer manuellement.

On examine l'influence de la pression et de la température (deux facteurs) sur le rendement y d'une réaction chimique (réponse). Le modèle choisi a priori est le suivant :

$$y = a_0 + a_1 \cdot x_1 + a_2 \cdot x_2 + a_{12} \cdot x_1 \cdot x_2$$

où x_1 et x_2 représentent respectivement les variables codées représentatives des facteurs pression et température. On utilise un tableau nommé **matrice d'expériences** pour récapituler l'ensemble des essais.

A partir du modèle, on peut écrire les 4 relations, en remplaçant les variables x_1 et x_2 par leurs valeurs (prises du tableau) dans chaque expérience.



N° Essai	Pression	Température	Rendement
1	-1	-1	60
2	1	-1	78
3	-1	1	63
4	1	1	89
Niveau -1	2 bars	50 °C	
Niveau +1	4 bars	70 °C	

Plans factoriels complet à deux niveaux 2^k

Exemple de cas d'étude (Suite):

A partir du modèle, on peut écrire les 4 relations, en remplaçant les variables x_1 et x_2 par leurs valeurs (prises du tableau) dans chaque expérience.

$$a_0 - a_1 - a_2 + a_{12} = 60$$

$$a_0 + a_1 - a_2 + a_{12} = 78$$

$$a_0 - a_1 + a_2 + a_{12} = 63$$

$$a_0 + a_1 + a_2 + a_{12} = 89$$

Après résolution, les coefficients sont égaux respectivement à :

$$a_0 = 72.5, a_1 = 11, a_2 = 3.5 \text{ et } a_{12} = 2$$

Le modèle mathématique s'écrit alors sous la forme (y représente le rendement) :

$$y = 72,5 + 11 \cdot x_1 + 3,5 \cdot x_2 + 2 \cdot x_1 \cdot x_2$$

où x_1 et x_2 sont les variables pression et températures exprimées en variables centrées réduites.

Discussion des résultats :

La variance σ^2 des variables Y_j étant supposée connue et identique sur le domaine d'étude (erreur aléatoire constante pour toutes les réponses Y_j) on en déduit que:

$$\text{var}(a_i) = N \cdot \text{var}(Y) / N^2 = \text{var}(Y) / N$$

On admet que chaque coefficient appartient à une population normale.

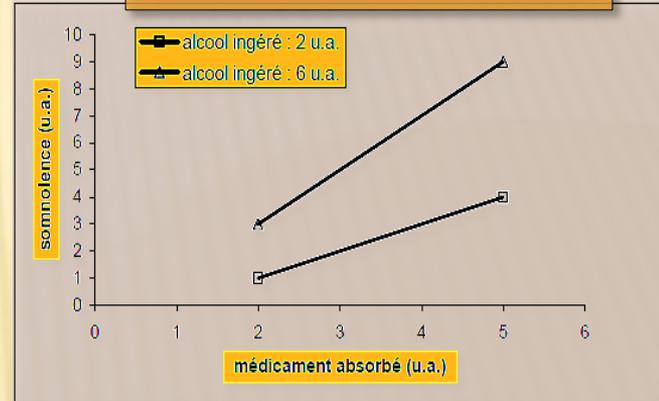
Avec un risque de 5%, si la variance estimée des mesures σ^2 est connue, une estimation ponctuelle de l'écart-type de la variable aléatoire a_i est $s(a_i)$:

$$s(a_i) = \sigma / \sqrt{N}$$

L'intervalle de confiance de tout coefficient est alors : $a_i \pm t_{0.975} \cdot (\sigma / \sqrt{N})$

où t est la variable de **Student** avec le nombre de degrés de liberté utilisé pour la détermination de s . A l'aide d'outils logiciels (comme **MINITAB**) on peut tracer la courbe ISO-REPOSE en transformant le modèle dans le domaine réel (naturel).

Interaction somnolence/médicaments



N° Essai	Pression	Température	Rendement
1	-1	-1	60
2	1	-1	78
3	-1	1	63
4	1	1	89
Niveau -1	2 bars	50 °C	
Niveau +1	4 bars	70 °C	

Plans factoriels complet à deux niveaux 2^k

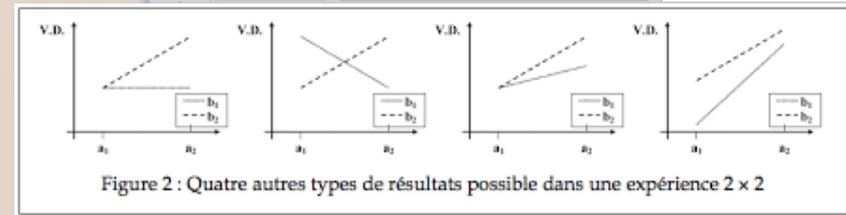
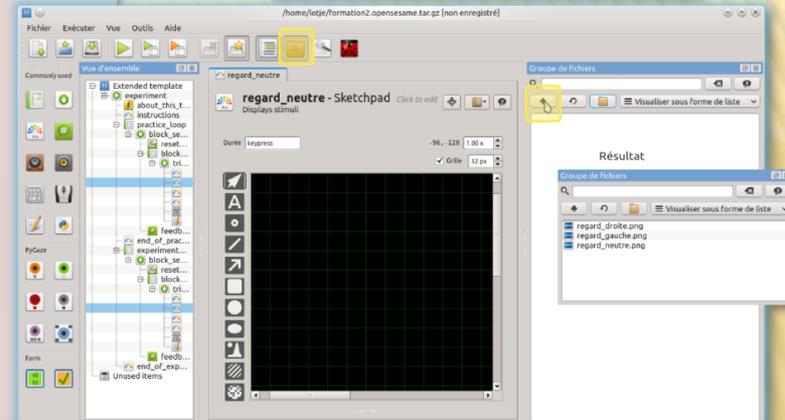
Logiciel OpenSeasame

Calcul des Effets et des Interactions

$E = (1/n) * X^T * Y$ où n est le nombre d'expériences.

Cas de deux facteurs à deux niveaux :

N°	0	1	2	12	y
1	+1	-1	-1	+1	y_1
2	+1	+1	-1	-1	y_2
3	+1	-1	+1	-1	y_3
4	+1	+1	+1	+1	y_4
Diviseur 2^k	4	4	4	4	
Effet	-		-	-	



Moyenne, colonne 0 : que des (+1)

Facteurs_(k), colonne i : alternance de blocs de 2^{i-1} "−1" suivis de 2^{i-1} "+1".

Interactions ($2^k - k - 1$), colonne ij pour une ligne ; produit des éléments des colonnes des facteurs i, j, ... concernées par l'interaction. Le diviseur : 2^k .

$$e_1 = (1/2^k) * [(+1)*y_1 + (+1)*y_2 + (+1)*y_3 + (+1)*y_4]$$

Plans factoriels complet à deux niveaux 2^k

Exercice :

Plan d'expérience avec 4 facteurs; $A\{a_1, a_2, a_3\}$; $B\{b_1, b_2\}$; $C\{c_1, c_2, c_3, c_4\}$; $D\{d_1, d_2\}$.

X représente le facteur et x_i est les modalités relatives.

1. Quel est le nombre des différents cas potentiels à tester.
2. Ecrire l'expression de la fonction de régression permettant de calculer tous les cas de valeurs.
3. Pour quatre essais choisis et exécutés, écrire la matrice d'Hadamard correspondante.
4. Dessiner le graphique d'un plan à 3 facteurs estimés à deux modalités.

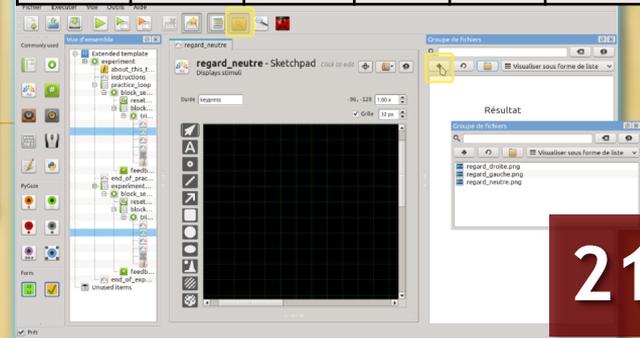
Exercice :

Un test d'arrachement lors de l'utilisation d'une colle met en jeu 3 facteurs : X_1 : la température de pressage. X_2 : la pression lors du pressage. X_3 : le temps de pressage. Les contraintes de l'expérimentation permettent de faire varier chacun des trois facteurs dans des fourchettes donnée. La réponse Y étudiée est la **tenue de la colle à l'arrachement** mesurée à l'aide d'un dynamomètre électronique. On admet un modèle polynomiale, linéaire par rapport aux coefficients. On effectue donc un plan d'expérience 2^3 , à l'aide d'une matrice factorielle. Les résultats sont dans le tableau ci-contre.

Le modèle s'écrit donc :

$$Y = 17,29 - 0,49 * X_1 + 0,01 * X_2 + 0,24 * X_3$$

Exp.	Moy.	x1	x2	x3	Y _{exp}
1	+1	-1	-1	-1	18.1
2	+1	+1	-1	-1	16.0
3	+1	-1	+1	-1	17.1
4	+1	+1	+1	-1	17.0
5	+1	-1	-1	+1	17.8
6	+1	+1	-1	+1	17.2
7	+1	-1	+1	+1	18.1
8	+1	+1	+1	+1	17.0
Div. 2 ³	8	8	8	8	
Effets	17.29	-0.49	0.01	0.24	



Plans factoriels complet à deux niveaux 2^k

Exercice :

Un industriel cherche à augmenter le rendement de sa fabrication. Il prépare un médicament à partir de plantes naturelles et cherche à améliorer le rendement d'extraction du principe actif. L'extraction est effectuée en présence de chlorure de sodium dont la concentration est de 50 grammes par litre et à une température de 70°C . L'industriel décide d'étudier ces deux facteurs et de les faire varier autour des consignes normales de fonctionnement. D'où les facteurs et le domaine d'étude :

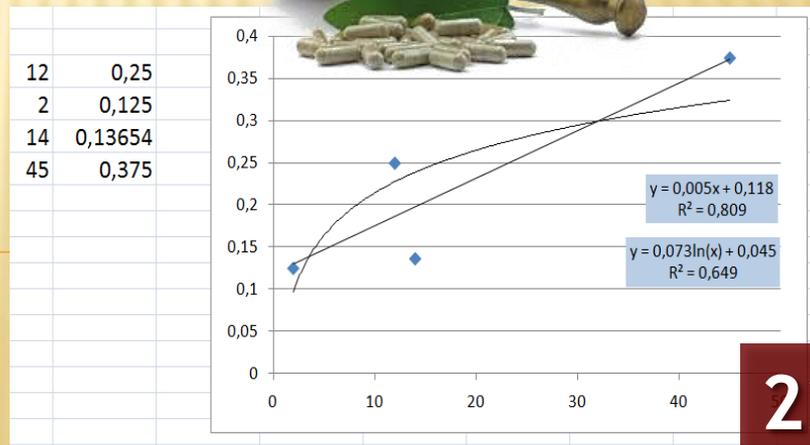
- Facteur 1 : **concentration** en chlorure de sodium entre 40 et 60 grammes.
- Facteur 2 : **température** entre 60°C et 80°C .

Formaliser ce problème (Donner le tableau et le plan de l'expérience ainsi que les équations et matrices correspondantes.)

Exercice :

Pour le tableau des données suivantes, ajouter un graphe en nuage de points, sous Microsoft Excel, en insérant deux estimations (régressions) avec leurs facteurs de qualité R^2 (Estimation du taux d'erreur au sens des moindres carrés).

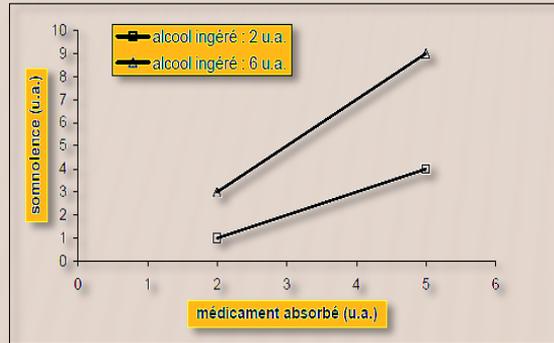
Expliquer le sens d'une régression si l'approximation est faite par un polynôme (cas linéaire) ou logarithmique (cas non linéaire).



Plans factoriels complets à deux niveaux 2^k

Test d'appréhension :

Des études sur la somnolence ont montré l'existence d'une interaction entre la quantité d'alcool ingéré X_A et la quantité d'un médicament particulier absorbé X_B . Le graphique suivant montre les valeurs de la somnolence mesurée en unité arbitraire de 0 à 10 en fonction des deux autres facteurs mesurés également avec des unités arbitraires.



Pour un plan d'expérimentation complet, déterminer :

- Le nombre d'essais à effectuer en laboratoire
- Donner le tableau du plan d'expérience.
- Donner l'expression de la régression simulant le phénomène y .
- Calculer les effets et les interactions.

Plans factoriels fractionnaires à deux niveaux 2^{k-p}

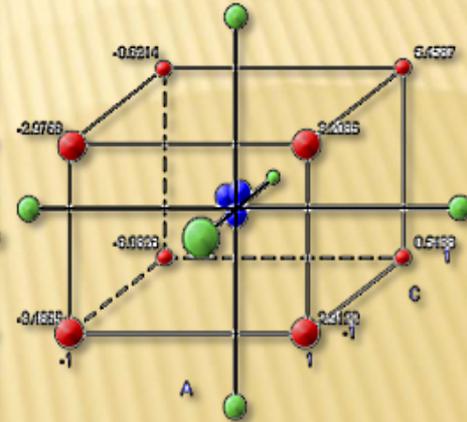
Définition :

Les plans factoriels fractionnaires sont des plans factoriels qui permettent d'étudier tous les facteurs mais dont le nombre d'essais est réduit par rapport aux plans factoriels complets. Un plan factoriel fractionnaire à 2^p fois moins d'essais que le factoriel complet correspondant.

A la fin d'un plan factoriel fractionnaire, on a un système de n équations à p coefficients inconnus avec p plus grand que n . On ne sait pas résoudre un tel système. Comme on ne peut pas augmenter le nombre d'équations, il faut diminuer le nombre d'inconnues. On y arrive en utilisant un artifice : on regroupe les coefficients de telle manière qu'il y ait n inconnues. On résout donc un système de n équations à n groupes de coefficients. On appelle ces groupes de coefficients, des **contrastes** ou des **aliases** et on dit que les coefficients sont **alliés dans les contrastes**.

Un plan 2^{k-p} permet d'étudier k facteurs prenant chacun deux niveaux. Le plan complet a été divisé par 2^p . Le nombre d'essais, dans ce cas est 2^{k-p} . Plus le nombre p augmente, plus la charge expérimentale va diminuer mais au détriment d'un risque de plus en plus grand sur la qualité des informations tirées du plan. Il faudra donc évaluer les risques avant de démarrer l'expérimentation et les minimiser en construisant le plan fractionnaire adéquat.

Courbes d'évolution d'un phénomène physique



Plans factoriels fractionnaires à deux niveaux 2^{k-p}

La difficulté, pour les plan fractionnaire est le nombre réel des facteurs à prendre en considération, oblige autant d'essais possibles afin d'équivaloir le nombre d'inconnues (coefficients) avec le nombre des équations qui les régissent.

Pour un plan complet équivalent, nous avons pour le cas de trois (03) facteurs 2^3 essais, d'où 08 équations de type $Y=X.a$.

$$Y = a_0 + a_1.x_1 + a_2.x_2 + a_3.x_3 + a_{12}.x_1.x_2 + a_{13}.x_1.x_3 + a_{23}.x_2.x_3 + a_{123}.x_1.x_2.x_3$$

Mais pour un cas de plan fractionnaire 2^{3-1} , la moitié des essais est exécutée réellement, donc moins d'équation pour pouvoir tirer les 8 inconnues. L'astuce est de choisir la transformation suivante $Y=X_s.l$ et dont le vecteur l est la substitution du vecteur a . On choisira les l coefficients (contrastes) de telle sorte que la matrice X_s soit une matrice orthogonale d'Hadamard.

A remarquer qu'un plan 2^3 est la juxtaposition de deux plan 2^{3-1} et 2^1 .

$$Y = l_0 + l_1.x_1 + l_2.x_2 + l_3.x_3$$

Les relations qui existent entre les contrastes inconnus et les vrais coefficients sont calculées de la manière suivante :

$$l_0 = a_0 + a_{123}$$

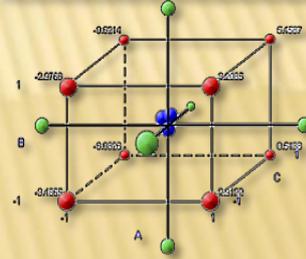
$$l_1 = a_1 + a_{23}$$

$$l_2 = a_2 + a_{13}$$

$$l_3 = a_3 + a_{12}$$

Les contrastes peuvent être aisément calculer mais reste à les expliquer et ainsi interpréter leur relation mathématique avec les coefficients.

Courbes d'évolution
d'un phénomène
physique



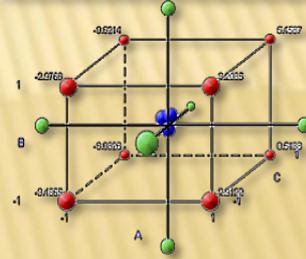
Plans factoriels fractionnaires à deux niveaux 2^{k-p}

L'interprétation des contrastes

Il est difficile de prédire les effets des facteurs sur eux-mêmes sous forme d'interactions double, triple ou supérieure, pour cela, on admet les hypothèses suivantes

- Hypothèse H1 : Les interactions d'ordre 3 (interaction entre 3 facteurs) ou d'ordre plus élevé sont considérées comme négligeables. On élimine ainsi un grand nombre d'inconnues. Mais attention cette hypothèse peut parfois être mise en défaut.
- Hypothèse H2 : Si un contraste est nul, cela peut signifier :
 - que les effets et les interactions aliasés sont tous nuls. Cette hypothèse est la plus probable et c'est celle que nous retiendrons sous le nom H2.
 - que les effets et les interactions aliasés se compensent. Cette hypothèse est peu probable et nous ne la retiendrons pas.
- Hypothèse H3 : Si deux contrastes sont faibles, on supposera que leur interaction l'est aussi. Si un contraste est faible et l'autre fort, on supposera que leur interaction est faible.
- Hypothèse H4 : Si deux contrastes sont forts, on se méfiera de leur interaction qui peut l'être également.

Courbes d'évolution
d'un phénomène
physique



Plans factoriels fractionnaires à deux niveaux 2^{k-p}

L'interprétation des contrastes

Pour un plan 2^3 avec 3 facteurs X_A , X_B et X_C , si on se contente de vouloir estimer l'influence de A, B et C sans se préoccuper des interactions, on a 4 coefficients à connaître : la moyenne et les effets des 3 facteurs, soit 4 essais.

$$y = a_0 + a_1 \cdot X_A + a_2 \cdot X_B + a_3 \cdot X_C$$

Un plan fractionnaire 2^{3-1} avec 4 essais suffira : 4 essais seront donc économisés. Néanmoins on n'obtiendra pas d'informations sur les interactions (AB, AC, BC). Si jamais une des interactions n'était pas négligeable, les coefficients du modèle seraient entachés d'erreur et le modèle ne conviendrait pas pour une projection ultérieure.

On conclusion, les plans fractionnaires sont dans beaucoup de cas utilisés en tant que plans d'échantillonnage destinés à déterminer les facteurs les plus dominants sur la fonction globale du processus étudié sans pour autant étudier les interactions d'ordre 2 et supérieures. Ce qui trouve sa justification dans le cas d'un nombre de facteurs très élevé.

Pour le cas d'un plan factoriel fractionnaire 2^{5-2} , la modélisation en un plan complet comportera 32 coefficients inconnus. Dans le cas **incomplet** établi avec seulement 8 essais de telle manière que la matrice soit une matrice orthogonale d'Hadamard. Nous obtenons, en conséquence, le plan d'expériences d'une telle matrice en prenant, comme choix, les colonnes 1, 2, 3, 1'2 et 1'3 de la matrice de calcul d'un plan 2^3 (tableau ci contre). L'étude du facteur 4 est donnée par les signes de l'interaction 1'2 et celle du facteur 5 par les signes de l'interaction 1'3.

N°	1	2	3	12	13
1	-1	-1	-1	+1	+1
2	+1	-1	-1	-1	-1
3	-1	+1	-1	-1	+1
4	+1	+1	-1	+1	-1
5	-1	-1	+1	+1	-1
6	+1	-1	+1	-1	+1
7	-1	+1	+1	-1	-1
8	+1	+1	+1	+1	+1

Plans factoriels fractionnaires à deux niveaux 2^{k-p}

L'interprétation des contrastes (suite)

Pour le plan complet, on obtient un système de 8 équations à 32 inconnues qui s'écrit sous forme matricielle :

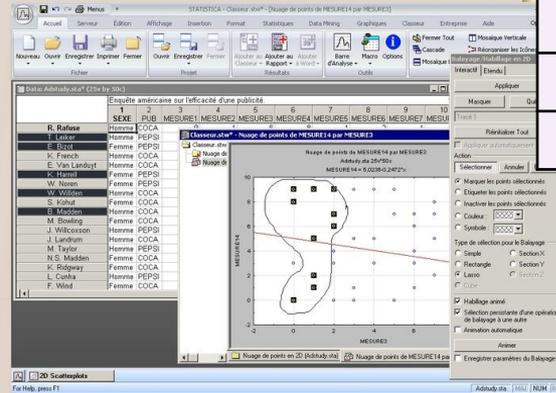
$$\begin{matrix} Y \\ (8 \times 1) \end{matrix} = \begin{matrix} X \\ (8 \times 32) \end{matrix} \cdot \begin{matrix} A \\ (32 \times 1) \end{matrix}$$

Pour réduire le nombre des inconnues, on introduit 8 contrastes.

$$\begin{matrix} Y \\ (8, 1) \end{matrix} = \begin{matrix} X_s \\ (8, 8) \end{matrix} \cdot \begin{matrix} A \\ (8, 1) \end{matrix}$$

La meilleure façon de savoir comment les coefficients sont aliasés dans les contrastes, est de faire appel aux logiciels. Dans les cas simples, on peut utiliser le calcul de Box. Pour cet exemple on trouve :

$$\begin{aligned} l_0 &= a_0 + a_{124} + a_{135} + a_{2345} \\ l_1 &= a_1 + a_{24} + a_{35} + a_{12345} \\ l_2 &= a_2 + a_{14} + a_{345} + a_{1235} \\ l_3 &= a_3 + a_{15} + a_{245} + a_{1234} \\ l_4 &= a_4 + a_{12} + a_{235} + a_{1345} \\ l_5 &= a_5 + a_{13} + a_{234} + a_{1245} \\ l_{23} &= a_{23} + a_{45} + a_{125} + a_{134} \\ l_{123} &= a_{123} + a_{25} + a_{34} + a_{145} \end{aligned}$$



N°	1	2	3	12	13
1	-1	-1	-1	+1	+1
2	+1	-1	-1	-1	-1
3	-1	+1	-1	-1	+1
4	+1	+1	-1	+1	-1
5	-1	-1	+1	+1	-1
6	+1	-1	+1	-1	+1
7	-1	+1	+1	-1	-1
8	+1	+1	+1	+1	+1

STATISTICA est un logiciel qui peut aider pour déterminer les regroupements des facteurs et/ou interactions sinon par la méthode de BOX.

Plans factoriels fractionnaires à deux niveaux 2^{k-p}

L'interprétation des contrastes (suite)

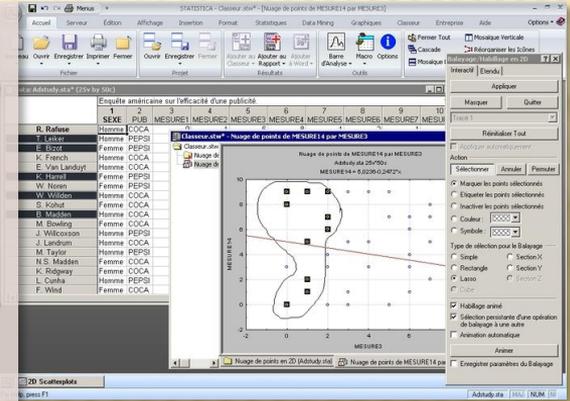
Le calcul de Box est un outil très utile pour l'écriture des contrastes. Cette méthode s'applique uniquement aux plans fractionnaires à 2 niveaux -1 et $+1$. Chaque colonne de la matrice des effets comprend uniquement des valeurs $+1$ ou -1 . On désigne par A la colonne de valeurs sous la variable A . On nomme I la colonne ne comportant que les valeurs $+1$.

Avec des colonnes quelconques A et B , les propriétés de la "multiplication des colonnes" sont les suivantes :

- $AB = BA$
- $AI = IA = A$ et $-IA = -A$
- $AA = I$

Une égalité entre deux colonnes A et B signifie que les termes des deux colonnes sont identiques.

La conception d'un plan fractionnaire 2^{3-1} permet en 4 expériences d'étudier 3 facteurs A , B et C et leurs interactions moyennant certains risques. Il faudra utiliser la matrice des effets d'un plan factoriel complet 2^2 et effectuer 4 expériences au lieu de 8.



N°	I	A	B	AB
1	+1	-1	-1	+1
2	+1	+1	-1	-1
3	+1	-1	+1	-1
4	+1	+1	+1	+1

Plans factoriels fractionnaires à deux niveaux 2^{k-p}

L'interprétation des contrastes (suite)

La matrice ci-contre ne comprend que 4 colonnes. Quatre autres termes sont à introduire : **C, AC, BC et ABC**. Pour cela on choisit comme aliase initial **C = AB**. Plutôt que d'écrire la transformation du polynôme, la méthode la plus rapide pour connaître tous les contrastes est de définir le **générateur d'aliases**.

L'écriture C = AB est équivalente à I = ABC

En effet : **C = AB \Rightarrow CC = ABC \Rightarrow I = ABC**

I = ABC est le générateur d'aliases. Les autres aliases sont obtenues en multipliant chaque colonne par le générateur d'aliases. D'après la propriété de multiplication avec la colonne I, on obtient alors :

I = ABC A = BC B = AC AB = C

Comme application : **A = AI = AABC = IBC = BC**

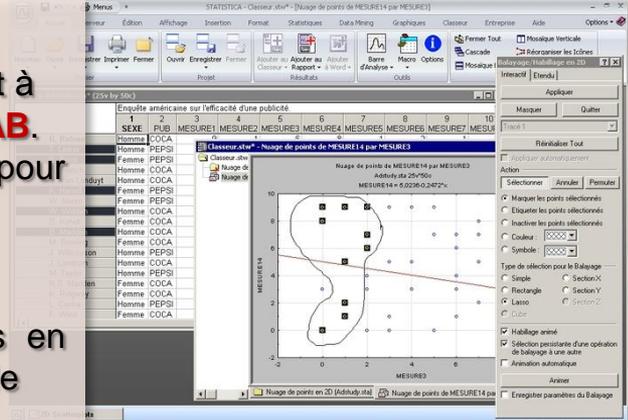
On en déduit alors la matrice des effets avec les contrastes :

Le contraste **h₂** se calcule par la méthode habituelle :

$$h_2 = (1/4) \cdot (-y_2 + y_3 - y_1 + y_4)$$

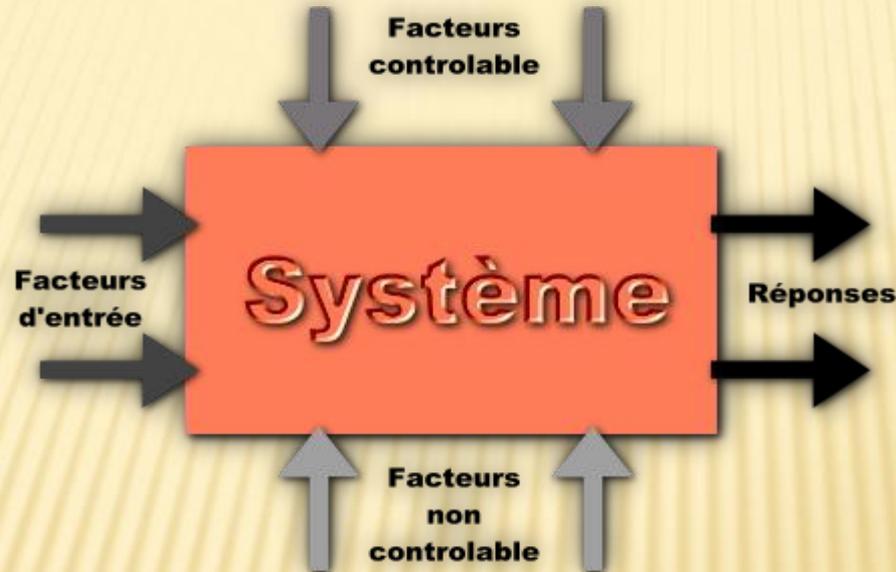
Puis on écrit la matrice d'expériences qui indique quels sont les niveaux à appliquer pour les facteurs **A, B et C**.

Les colonnes **C** et **AB** doivent être identiques.



N°	I	A	B	AB	Y
	= ABC	= BC	= AC	= C	
1	+1	-1	-1	+1	y1
2	+1	+1	-1	-1	y2
3	+1	-1	+1	-1	y3
4	+1	+1	+1	+1	y4
contraste	h1	h2	h3	h4	

A suivre ...



MERCI POUR VOTRE ATTENTION

Fin du onzième chapitre

Références

1. L.-V. Bertalanfy, 'General System Theory', Edition MASSON, 1972.
2. D. Collombier, 'Plans d'expérience factoriels : construction et propriétés des fractions de plans', éditions Springer, 1996.
3. F. Husson et J. Pagès, 'Statistiques générales pour utilisateurs', PUR, 2005.
4. Douglas C. Montgomery, 'Design and Analysis of Experiments', 7th Edition, éditions Wiley, 2009.
5. Open Sesame, <http://osdoc.cogsci.nl/>