LES PLANS FACTORIELS COMPLETS

Claude Hoinard

Laboratoire de Biophysique et Mathématiques

Faculté de Pharmacie de Tours

LES PLANS EXPERIMENTAUX

- ## Introduction
  - Une **expérience** est une intervention volontaire dans un système en fonctionnement pour mesurer les effets de cette intervention. Seule l'expérience est capable d'apporter des renseignements sur les relations de cause à effets.

Pourquoi un plan expérimental ?

L'expérimentation coûte cher, il faut avoir pour objectif d'obtenir les informations les plus fiables et les plus efficaces possibles en un minimum d'essais :

L'EXPÉRIMENTATION DOIT ÊTRE OPTIMISÉE

↓

Nécessité d'une recherche expérimentale bien planifiée

Notion de plan efficace

« Ne faire varier qu'un facteur à la fois »
= stratégie sécurisante mais résultats médiocres

Un ***plan*** d 'expériences ***efficace*** se propose donc de faire varier plusieurs facteurs à la fois selon des règles d 'organisation précises et rigoureuses : les plans que nous allons étudier sont **multi factoriels** et sont appelés habituellement ***plans factoriels***.

Un peu de vocabulaire . . .

- Un **facteur** (*quantitatif ou qualitatif*) est un état du système étudié dont la variation est susceptible d'en modifier le fonctionnement (ex : *température, concentration d'un réactif…*)
- La **réponse** du système correspond au paramètre mesuré ou observé pour connaître l'effet des facteurs étudiés sur le système (ex : *rendement d'une réaction…*)

Méthodologie dans l'étude d 'un phénomène

Lors de l 'étude d 'un phénomène, plusieurs questions se posent, auxquelles répondent différents types de plans. On peut distinguer 3 grandes étapes dans l 'acquisition des connaissances :

❶ Recherche des facteurs influents

❷ Modélisation

❸ Optimisation

⇒ la recherche des facteurs influents

Parmi tous les facteurs susceptibles d 'influer sur le phénomène (c 'est à dire sur la ou les réponses mesurées du phénomène),

- lesquels ont une influence significative ?
- que vaut quantitativement cette influence ?
- existe-t 'il des interactions entre facteurs ?

Les plans qui permettent de rechercher les facteurs influents sont les ***plans factoriels complets ou fractionnaires.***

⇒ la modélisation

Quand les facteurs influents ont été identifiés et leur importance quantifiée, on recherche ensuite l 'équation permettant de décrire les variations de la réponse étudiée en fonction de celles des facteurs influents ; cette seconde étape constitue la ***modélisation***.

Les plans factoriels suffisent parfois. Il se peut dans d 'autres cas que l 'on soit obligé de faire appel à des plans plus complexes tels que les **plans composites centrés**.

⇒ l 'optimisation

Pour finir, on cherche en général à déterminer quelles conditions expérimentales (les valeurs prises par les facteurs influents) permettent d 'obtenir le meilleur résultat pour la réponse. Cette étape porte le nom d '***optimisation***.

Il existe plusieurs méthodes pour ce faire ; citons,

→ l 'étude des courbes iso réponses

→ la méthode du simplex

La recherche des facteurs influents

préliminaire

Avant de débuter, plusieurs questions à se poser :

⇨ Quelles réponses choisir ?

Les divers aspects du phénomène étudié se traduisent par autant de réponses possibles ; les réponses choisies seront celles qui caractérisent le mieux les propriétés intéressant l'expérimentateur pour son application ; chaque réponse sera analysée d'abord seule, puis une synthèse globale sera faite ensuite en tenant compte des résultats obtenus avec toutes les réponses.

⇨Quels facteurs à étudier ?

Ils seront choisis en fonction de la connaissance que l 'on a déjà du phénomène. Leur nombre conditionne le nombre d 'essais à réaliser ; des contraintes économiques et de temps interviennent donc dans ce choix.

⇨Autres facteurs à considérer

On ne peut pas étudier tous les facteurs susceptibles d 'agir sur une réponse ... Il faudra cependant en faire une liste aussi exhaustive que possible et distinguer les **facteurs étudiés**, les **facteurs contrôlables mais non retenus dans l 'étude** (leur valeur devra être maintenue constante dans tous les essais du plan) et les **facteurs non contrôlables** (non maîtrisables ou non identifiés).

Les facteurs non contrôlables, dont les niveaux varient de façon anarchique d 'un essai à l 'autre lors de l 'expérimentation, **ont pour effet de faire fluctuer la réponse mesurée.**
On devra donc s 'assurer que ces fluctuations sont aléatoires et estimer cette dispersion des réponses afin d 'en tenir compte lors de l 'analyse des résultats.
Pour ce faire, il est fréquent d 'être obligé d 'effectuer des **essais complémentaires**.

⇨ Limites du domaine expérimental

Il faut également fixer les limites du domaine expérimental, c 'est à dire les valeurs extrêmes prises par les facteurs étudiés. Notons dès maintenant que ce sont ces valeurs limites qui sont utilisées lors des manipulations.

LES PLANS FACTORIELS COMPLETS A 2 NIVEAUX

- **présentation**
  - **n** facteurs (A,B,C,…) sont étudiés chacun à 2 niveaux (2 valeurs) A_0/A_1; B_0/B_1; C_0/C_1…) choisis par l'expérimentateur
  - Le plan est dit **factoriel complet** si tous les niveaux de tous les facteurs sont programmés pour varier simultanément et de manière équilibrée dans l'ensemble de l'expérimentation.

- **2^2** : 2 facteurs à 2 niveaux définissent 4 conditions expérimentales soit,
  - 4 essais s'il n'est pas prévu de répétitions
  - 8 essais si le plan est « dupliqué » (2 déterminations par condition)
  - 12 essais si le plan est « tripliqué »…..etc..

- **2^3** : 3 facteurs à 2 niveaux définissent 8 conditions expérimentales soit,
  - 8 essais s'il n'est pas prévu de répétitions, 16 si le plan est « dupliqué » etc…
- **2^6** : 6 facteurs à 2 niveaux ,soit 64 conditions expérimentales soit,
  - 64 essais (pas de répétition) , 128 (plan « dupliqué ») etc..

Ces « Plans 2 » sont ceux qu'il faut utiliser pour rechercher les facteurs agissant sur une réponse mesurée. Ce sont les plus simples à interpréter et ils présentent le meilleur rapport coût/efficacité.

PLAN D'ETUDE

Nous examinerons :

1. ce qu'on entend par EFFETS PRINCIPAUX et INTERACTIONS et comment on les évalue
2. à partir des résultats obtenus, comment on peut sélectionner les facteurs qui ont une influence réelle sur la réponse et ceux qui n'ont pas d'action.
3. nous examinerons enfin le modèle associé à ces « plans 2 », passage indispensable pour interpréter en profondeur les résultats obtenus et préalable à l'OPTIMISATION.

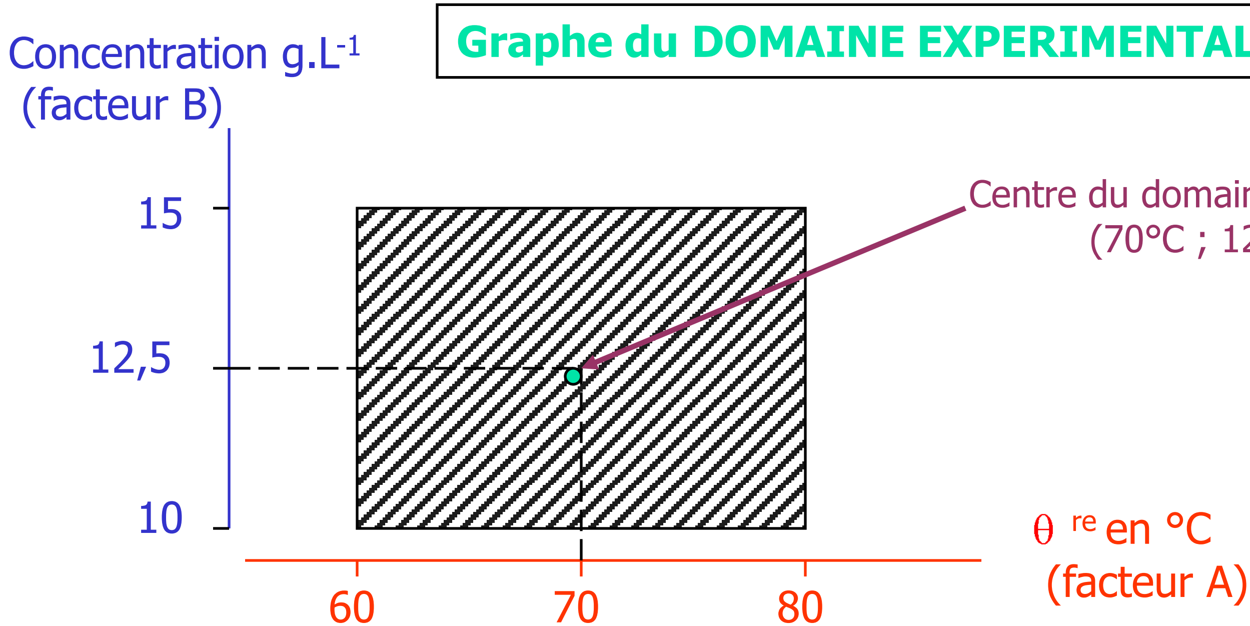
1- DEFINITION DES EFFETS PRINCIPAUX ET DES INTERACTIONS - CALCULS

1-1 : Le plan 2^2

Exercice 1 : étude de l'influence de la température et de la concentration C d'un réactif sur le rendement y (en %) d'une réaction chimique.

Il a été décidé d'expérimenter la θ re entre 60°C et 80°C et la concentration entre 10 $g.L^{-1}$ et 15 $g.L^{-1}$ en se limitant à 2 valeurs par facteur.

Graphe du DOMAINE EXPERIMENTAL (hachuré)



Avec 4 conditions expérimentales, le maximum d'efficacité dans l'étude est obtenu lorsque celles-ci sont situées aux extrémités du domaine expérimental.

Soit	A_0B_0	:	60°C / 10 $g.L^{-1}$
	A_1B_0	:	80°C / 10 $g.L^{-1}$
	A_0B_1	:	60°C / 15 $g.L^{-1}$
	A_1B_1	:	80°C / 15 $g.L^{-1}$

⇨ Cette propriété, maximum efficacité = extrémités du domaine est générale à tous les essais factoriels 2.

1-1-1 : la matrice d'expérience

Il est commode de symboliser par –1 le niveau BAS de chaque facteur et par +1 le niveau HAUT. Ce qui permet de rassembler les éléments relatifs à chaque facteur dans un tableau appelé **MATRICE D'EXPERIENCE** que l'on présente en correspondance avec une colonne donnant les résultats expérimentaux de la la ***réponse y***.

niveau -1	60°C	10 $g.L^{-1}$
niveau +1	80°C	15 $g.L^{-1}$

n° essai	θ re facteur A	Concentration facteur B	Réponse rendement %
1	-1	-1	y_1 = 60
2	+1	-1	y_2 = 70
3	-1	+1	y_3 = 80
4	+1	+1	y_4 = 90

La colonne n°essai, repère de la condition expérimentale (1=A_0B_0, 2= A_1B_0, 3= A_0B_1, 4= A_1B_1) permet d'identifier la réponse : y_1 est la réponse pour la condition A_0B_0….etc…

❑ Les facteurs ainsi codés sont dits **FACTEURS CENTRÉS RÉDUITS** et notés $\mathbf{x_A}$:

- Dans le cas de la température (facteur A)

  si 60°C correspond à –1

  et 80°C correspond à +1

cela signifie que 70°C correspond à $x_A=0$ et qu'entre x_A et θ température il existe la relation

$$x_A = \frac{\theta - 70}{10}$$

65°C correspond donc à $x_A=-0,5$ et 78°C à $x_A=+0,8$

- Pour la concentration, $x_B=-1$ pour 10 g.L^{-1} et $x_B=+1$ pour 15 g.L^{-1}

  $x_B=0$ correspond à 12,5 g.L^{-1} et on a :

$$x_B = \frac{C - 12,5}{2,5}$$

→ On notera que le centre du domaine expérimental correspond au niveau 0 des facteurs centrés réduits.

1-1-2 : effet global et effet moyen d'un facteur. Représentation graphique

- **EFFET GLOBAL**

C'est la variation de réponse y quand le facteur passe du niveau BAS (-1) au niveau HAUT (+1)

θ re A	Conc B	Réponse %
-1	-1	$y_1 = 60$
+1	-1	$y_2 = 70$
-1	+1	$y_3 = 80$
+1	+1	$y_4 = 90$

Par suite de l'organisation des expériences factorielles 2^2, un effet peut s'évaluer de 2 façons :

- température (A)

à 10 g.L^{-1} effet global = y_2-y_1 =70 – 60 = 10%

à 15 g.L^{-1} effet global = y_4-y_3 =90 – 80 = 10%

- concentration (B)

à 60°C effet global = y_3-y_1 =80 – 60 = 20%

à 80°C effet global = y_4-y_2 =90 – 70 = 20%

- Dans cet exemple particulier, l'effet d'un facteur est exactement le même quelque soit le niveau de l'autre : les 2 facteurs agissent indépendamment. On dit aussi qu'il n'y a pas d'interaction entre les 2 facteurs.

- Notons qu'en général, même lorsqu'il y a indépendance des facteurs, les effets calculés ne donneront pas exactement la même valeur pour les 2 niveaux de l'autre facteur à cause des erreurs aléatoires de mesures des réponses ; on en fera la moyenne :

  - effet global de la température : $\frac{1}{2}[(y_2-y_1)+(y_4-y_3)]$
  - effet global de la concentration : $\frac{1}{2}[(y_3-y_1)+(y_4-y_2)]$

EFFET MOYEN

- L'effet global correspond à une variation du facteur centré réduit de 2 unités (de –1 à +1)
- L'effet moyen correspond à une variation de 1 unité du facteur centré réduit (de 0 à 1 par exemple)

$\boxed{\text{L'effet moyen noté } \mathbf{E} \text{ est la moitié de l'effet global}}$

C'est l'effet moyen que nous utiliserons. On a donc :

- effet moyen de la température : $\frac{1}{4}(-y_1 + y_2 - y_3 + y_4)$
- effet moyen de la concentration : $\frac{1}{4}(-y_1 - y_2 + y_3 + y_4)$

Dans l'exemple, on a donc $E_A = 5\ \%$ (θ re) et $E_B = 10\ \%$ (concentration)

- Remarque : le calcul de la moyenne générale des essais s'écrit

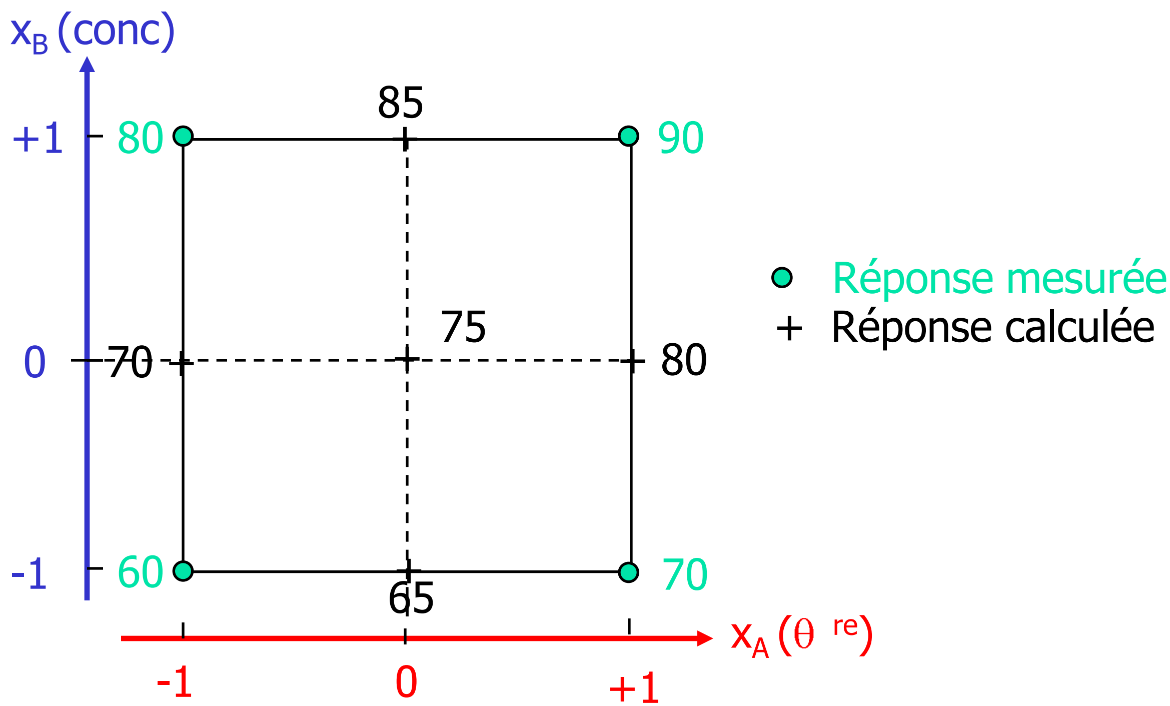
$$\bar{y}_c = \frac{1}{4}(y_1 + y_2 + y_3 + y_4)$$

Elle vaut donc dans l'exemple 75% et elle peut être considérée comme la valeur calculée de la réponse quand les facteurs centrés réduits sont au niveau 0, c'est à dire au centre du domaine expérimental.

REPRESENTATIONS GRAPHIQUES

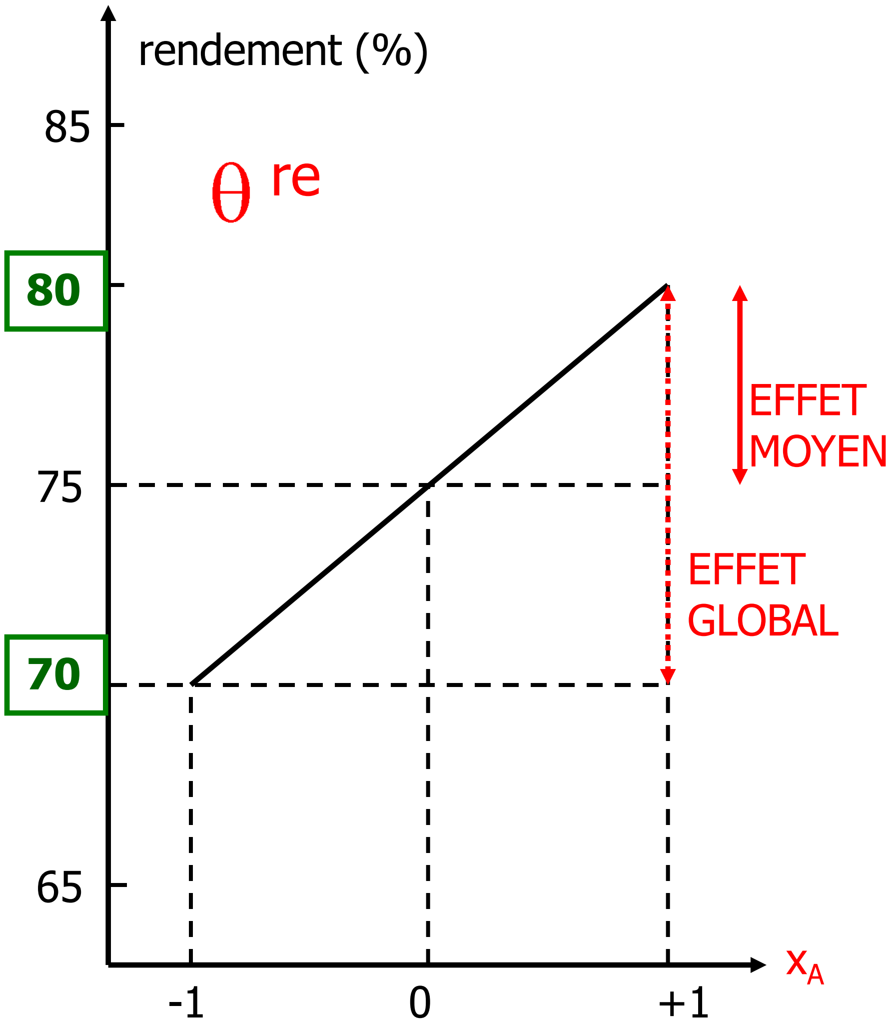
- Pour visualiser les résultats on peut

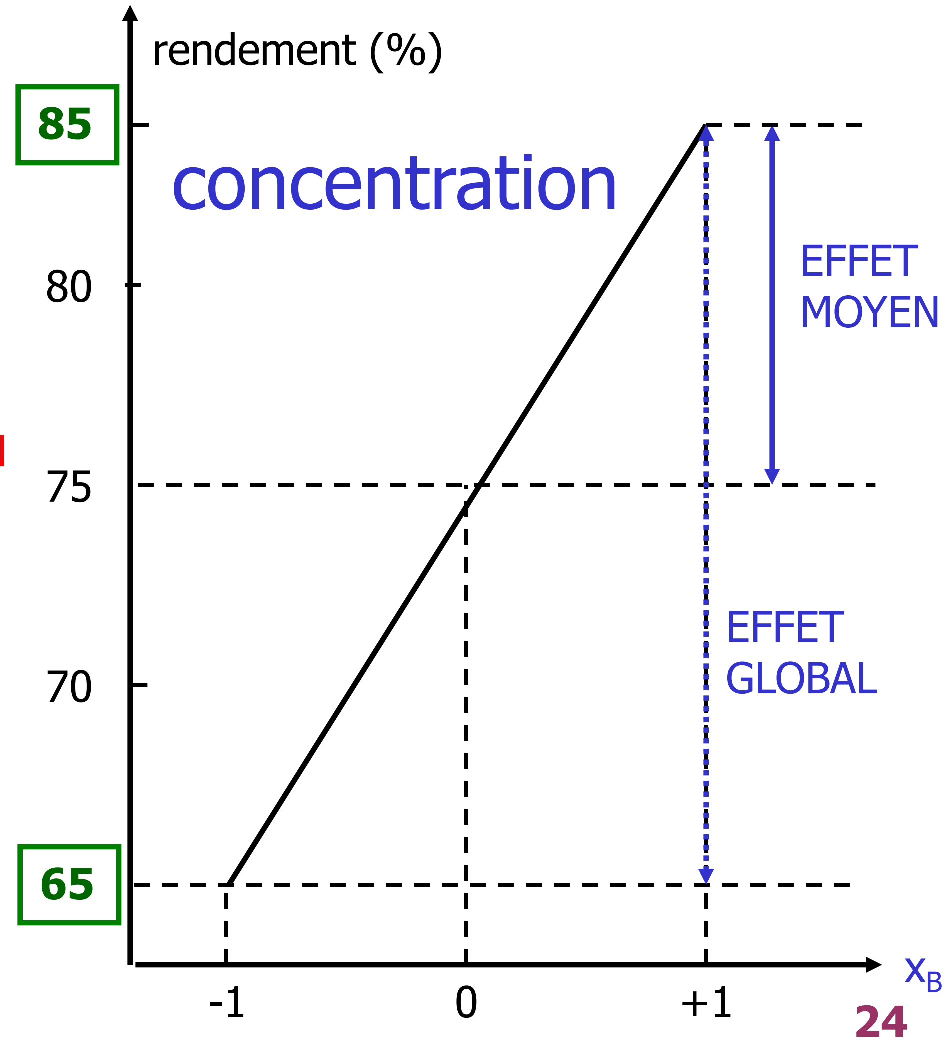
1. reporter les réponses mesurées et calculées sur le domaine expérimental



REPRESENTATIONS GRAPHIQUES

2. reporter les réponses moyennes calculées en fonction du facteur réduit pour le niveau 0 de l’autre, ce qui permet de faire apparaître l’effet de ce facteur





1-1-3 : interaction entre 2 facteurs. Évaluation

- Exercice 2 **:** même étude que l'exercice 1 réalisée cette fois en présence d'un catalyseur ; les conditions expérimentales sont identiques.Les résultats de rendement sont :

n° essai	1	2	3	4
Rendement (%)	60	70	80	95

La moyenne calculée des essais est devenue : $\overline{y}_c = \frac{1}{4}(60+70+80+95) = 76,25$

La présence du catalyseur est favorable. Qu'a t-il changé dans les effets ?

- effet de la température :

à 10 g.L^{-1} : $E_\theta = \frac{1}{2}(70-60) = 5\%$ à 15 g.L^{-1} : $E_\theta = \frac{1}{2}(95-80) = 7,5\%$

- effet de la concentration :

à 60°C : $E_{conc} = \frac{1}{2}(80-60) = 10\%$ à 80°C : $E_{conc} = \frac{1}{2}(95-70) = 12,5\%$

L'EFFET D'UN FACTEUR DEPEND DU NIVEAU DE L'AUTRE : ON DIT QU'IL Y A **INTERACTION** ENTRE LES FACTEURS θ ^{re} et CONCENTRATION

EVALUATION DE L'INTERACTION (E_{AB})

> Elle se définit en grandeur comme la moitié de la différence entre l'effet moyen d'un facteur au niveau HAUT de l'autre et de l'effet moyen de ce facteur au niveau BAS de l'autre

- Appliquons à l'exemple 2 :

E(θ re X conc) = $\frac{1}{2}(7,5-5)=1,25\%$ ou $\frac{1}{2}(12,5-10)=1,25\%$

- D'une manière générale, l'effet de l'interaction s'exprime en fonction des réponses individuelles par :

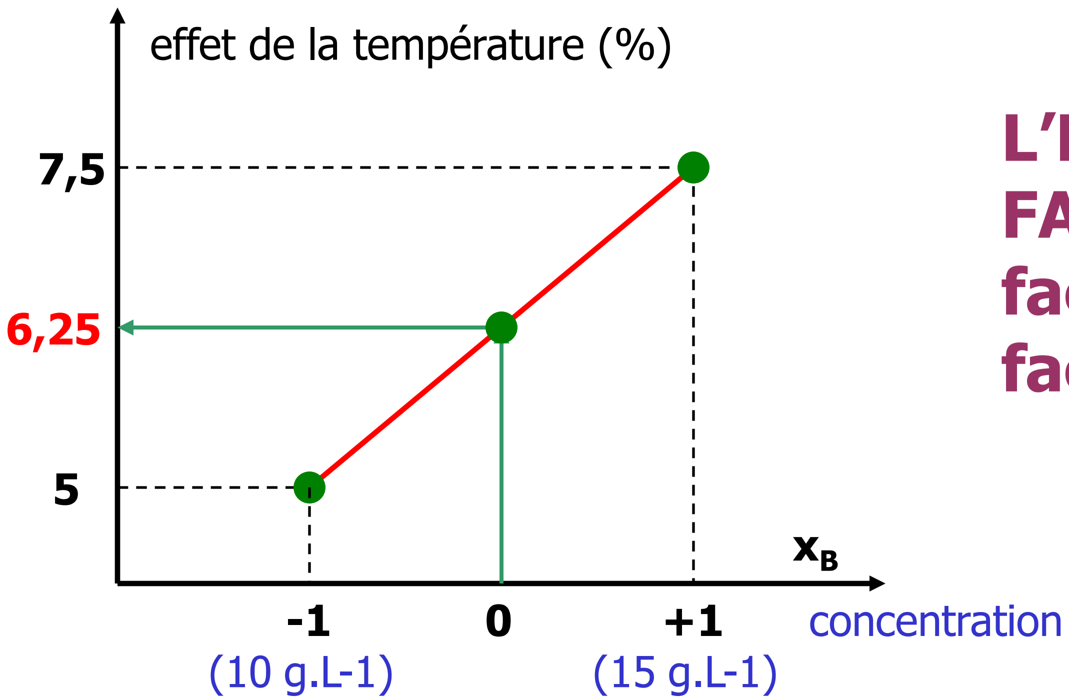
$$E_{AB}=\frac{1}{4}(y_1-y_2-y_3+y_4)$$

▪ SIGNIFICATION DE L'EFFET D'UN FACTEUR QUAND IL Y A INTERACTION

On peut calculer E_A et E_B grâce aux formules vues précédemment. Dans l'exemple 2 on trouve :

- température : $E_A = 6{,}25$ % - concentration : $E_B = 11{,}25$ %

La signification de l'effet d'un facteur, appelé EFFET PRINCIPAL, apparaît quand on représente la variation de l'effet en fonction du niveau de l'autre facteur.



L'EFFET PRINCIPAL D'UN FACTEUR est l'effet de ce facteur quand l'autre facteur est au niveau 0

1-1-4 : matrice des effets ; pratique des calculs

L'ensemble des résultats à partir des réponses mesurées est donné par les formules :

$$\begin{cases} E_A = \dfrac{1}{4}(-y_1 + y_2 - y_3 + y_4) & \text{effet principal de A} \\[2ex] E_B = \dfrac{1}{4}(-y_1 - y_2 + y_3 + y_4) & \text{effet principal de B} \\[2ex] E_{AB} = \dfrac{1}{4}(+y_1 - y_2 - y_3 + y_4) & \text{effet interaction AB} \\[2ex] \overline{y}_c = \dfrac{1}{4}(+y_1 + y_2 + y_3 + y_4) & \text{réponse calculée au centre du domaine} \end{cases}$$

Il est commode de dresser le tableau suivant appelé

MATRICE DES EFFETS

n° essai	A	B	AB	M
1	-1	-1	+1	+1
2	+1	-1	-1	+1
3	-1	+1	-1	+1
4	+1	+1	+1	+1

réponse
y_1
y_2
y_3
y_4

- La colonne **AB** matérialisant l'interaction s'obtient en multipliant ligne par ligne les colonnes A et B
- La colonne **M** (réponse au centre du domaine) ne contient que des +1

Pour calculer les effets, il suffit de multiplier chaque colonne de la matrice par la colonne réponse , ligne par ligne , d'additionner les produits et de diviser par 4.

Ces calculs sont très faciles à effectuer avec un **TABLEUR**, EXCEL par exemple (utilisation de la fonction SOMMEPROD)

1-2 : Le plan 2^3

Exercice 3 : dans une solution pharmaceutique habituellement fabriquée à 30°C, sous agitation (200 tours.min^{-1}) un léger trouble apparaît.

L'expérimentateur désire connaître la (ou les) cause(s) et pense que 3 facteurs peuvent jouer :

- la température
- la vitesse d'agitation
- la concentration d'un additif présent à 0,30 %

Le trouble se mesure par un indice d'opacité traduisant l'impression visuelle que donne l'intensité du louche.

- Il est décidé d'organiser une expérience factorielle 2^3

facteurs	niveaux
A = θ re	20°C – 40°C
B = vitesse agitation	100 $t.min^{-1}$ - 300 $t.min^{-1}$
C = concentration additif	0,1 % - 0,5 %

- Il y a donc 8 conditions expérimentales

$1 = A_0B_0C_0$ $\quad$ $5 = A_0B_0C_1$

$2 = A_1B_0C_0$ $\quad$ $6 = A_1B_0C_1$

$3 = A_0B_1C_0$ $\quad$ $7 = A_0B_1C_1$

$4 = A_1B_1C_0$ $\quad$ $8 = A_1B_1C_1$

1-2-1 : domaine expérimental – matrice d'expérience

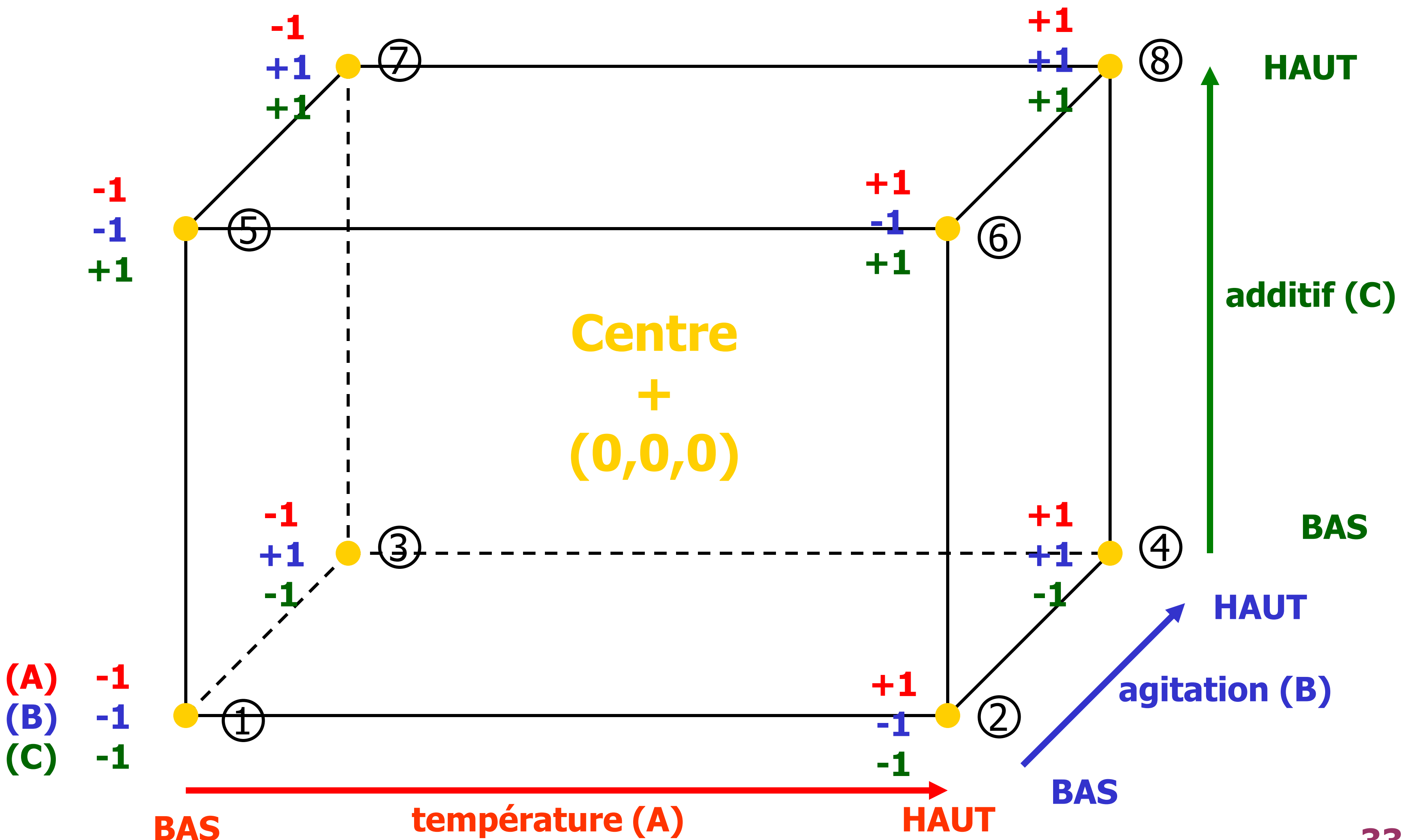
Codons comme précédemment les niveaux des facteurs (-1, +1) en utilisant les facteurs centrés réduits.

2 remarques :

1. Le **centre du domaine** (0, 0, 0) correspond à la condition expérimentale habituellement utilisée (30°C – 200 t.min^{-1} – 0,3 %)

2. Comme pour un plan 2^2, les conditions expérimentales se trouvent aux **extrémités du domaine**. C'est avec ces conditions que l'expérimentation sera la plus efficace.

Plan 2^3 : Le domaine expérimental peut être représenté par un cube.



Plan 2^3 : **MATRICE D'EXPERIENCE**

niveau -1	20°C	100 $t.min^{-1}$	0,1 %
niveau +1	40°C	300 $t.min^{-1}$	0,5 %
n° essai	A (θ re)	B (agit.)	C (add.)
1	-1	-1	-1
2	+1	-1	-1
3	-1	+1	-1
4	+1	+1	-1
5	-1	-1	+1
6	+1	-1	+1
7	-1	+1	+1
8	+1	+1	+1

réponse
y_1
y_2
y_3
y_4
y_5
y_6
y_7
y_8

IMPORTANT : noter l'organisation des signes :

- pour A : succession 1- puis 1+
- pour B : succession 2- puis 2+
- pour C : succession 4- puis 4+

1-2-2 : Plan 2^3 - effets des facteurs principaux et interactions.

- **FACTEURS PRINCIPAUX**

Avec ce type de plan, l'effet d'un facteur est évalué de 4 façons.

Par exemple pour A : y_2-y_1 , y_4-y_3 , y_6-y_5 , y_8-y_7

A (θ^{re})	réponse
-1	y_1
+1	y_2
-1	y_3
+1	y_4
-1	y_5
+1	y_6
-1	y_7
+1	y_8

- L'effet moyen de **A** s'écrit :

$$E_A = \frac{1}{8}(-y_1 + y_2 - y_3 + y_4 - y_5 + y_6 - y_7 + y_8)$$

On retrouve l'organisation des signes de la matrice d'expérience

- De même pour les facteurs **B** et **C**

$$E_B = \frac{1}{8}(-y_1 - y_2 + y_3 + y_4 - y_5 - y_6 + y_7 + y_8)$$

$$E_C = \frac{1}{8}(-y_1 - y_2 - y_3 - y_4 + y_5 + y_6 + y_7 + y_8)$$

• **INTERACTIONS DE 2EME ORDRE**

• Avec 3 facteurs étudiés, il y a 3 interactions possibles mettant en jeu 2 facteurs : **A.B** **A.C** **B.C**

Elles sont appelées interactions de 2ème ordre. Leur nombre s'obtient en combinant les 3 facteurs 2 à 2 : $C_3^2 = 3$

• La grandeur de ces interactions se définit comme pour une factorielle 2^2 : moitié de la différence entre l'effet d'un facteur au niveau HAUT de l'autre et l'effet de ce facteur au niveau BAS de l'autre.

• Considérons par exemple l'interaction A.B en considérant l'effet de A :

A	B	rép
-1	-1	y_1
+1	-1	y_2
-1	+1	y_3
+1	+1	y_4
-1	-1	y_5
+1	-1	y_6
-1	+1	y_7
+1	+1	y_8

- y_4-y_3 et y_8-y_7 effets de A au niveau HAUT de B
- y_2-y_1 et y_6-y_5 effets de A au niveau BAS de B

On a donc :

$$E_{AB} = \frac{1}{2}\left[\frac{1}{4}(y_4 - y_3 + y_8 - y_7) - \frac{1}{4}(y_2 - y_1 + y_6 - y_5)\right]$$

$$E_{AB} = \frac{1}{8}(+y_1 - y_2 - y_3 + y_4 + y_5 - y_6 - y_7 + y_8)$$

On pourrait calculer de la même façon E_{AC} et E_{BC}. Ce n'est pas nécessaire, la construction de la matrice des effets les donne automatiquement.

• INTERACTIONS DE 3EME ORDRE

• Il existe aussi une possibilité d'interaction entre les 3 facteurs pris ensemble, notée **A.B.C**, et qui porte le nom d'*interaction du 3ème ordre*

• Elle peut se définir de la façon suivante :

n°	A	B	C	rép
1	-1	-1	**-1**	y_1
2	+1	-1	**-1**	y_2
3	-1	+1	**-1**	y_3
4	+1	+1	**-1**	y_4
5	-1	-1	**+1**	y_5
6	+1	-1	**+1**	y_6
7	-1	+1	**+1**	y_7
8	+1	+1	**+1**	y_8

- Les 4 premières lignes de la matrice d'expérience correspondent à un plan 2^2 au niveau BAS du facteur C et :

$$(E_{AB})_{C_0} = \frac{1}{4}(y_1 - y_2 - y_3 + y_4)$$

- Les 4 dernières lignes correspondent à un plan 2^2 au niveau HAUT du facteur C

$$(E_{AB})_{C_1} = \frac{1}{4}(y_5 - y_6 - y_7 + y_8)$$

• si $(E_{AB})_{C_0} = (E_{AB})_{C_1}$, cela signifie que l'interaction AB est indépendante du niveau du 3e facteur : il n'y a pas d'interaction de 3ème ordre.

• si $(E_{AB})_{C_0} \neq (E_{AB})_{C_1}$, il y a interaction de 3ème ordre et on la définit en grandeur comme la moitié de la différence de l'interaction AB au niveau HAUT de C et de l'interaction AB au niveau BAS de C.

$$E_{ABC} = \frac{1}{2}\left[(E_{AB})_{C_1} - (E_{AB})_{C_0}\right] = \frac{1}{8}(-y_1 + y_2 + y_3 - y_4 + y_5 - y_6 - y_7 + y_8)$$

Cette expression est donnée automatiquement lors de la construction de la matrice des effets.

• INTERACTIONS : EN PRATIQUE

• On trouve souvent dans les applications des interactions de 2ème ordre – il est tout à fait concevable que 2 facteurs n’agissent pas indépendamment sur la réponse –

• En revanche, l’existence d’une interaction de 3ème ordre suppose un système complexe mettant en jeu les 3 facteurs simultanément et se rencontre très rarement.

1-2-3 : Plan 2^3 - matrice des effets - calculs

- Dans une expérience factorielle 2^3, il y a au total 8 éléments à calculer à partir des réponses :
  - les effets principaux des 3 facteurs
  - les effets des 3 interactions du 2ème ordre
  - l'effet de l'interaction du 3ème ordre
  - la réponse $\overline{y}_C$ au centre du domaine
- La matrice des effets comportera donc 8 colonnes (avec en plus la colonne n° essai et la colonne des réponses mesurées)

  - 3 colonnes sont déjà connues : celles des facteurs de la matrice d'expérience.
  - les 3 colonnes (AB), (AC), (BC) s'obtiennent en multipliant les colonnes A et B ligne par ligne, idem pour A et C et pour B et C.
  - la colonne ABC représentant l'interaction de 3ème ordre s'obtient en multipliant ligne par ligne les 3 colonnes A, B et C.
  - la colonne M pour calculer y_C est constituée de +1

Plan 2^3 - matrice des effets - calculs (suite)

n° essai	A	B	C	AB	AC	BC	ABC	M	rép
1	-1	-1	-1	*+1*	*+1*	*+1*	*-1*	*1*	y_1
2	+1	-1	-1	*-1*	*-1*	*+1*	*+1*	*1*	y_2
3	-1	+1	-1	*-1*	*+1*	*-1*	*+1*	*1*	y_3
4	+1	+1	-1	*+1*	*-1*	*-1*	*-1*	*1*	y_4
5	-1	-1	+1	*+1*	*-1*	*-1*	*+1*	*1*	y_5
6	+1	-1	+1	*-1*	*+1*	*-1*	*-1*	*1*	y_6
7	-1	+1	+1	*-1*	*-1*	*+1*	*-1*	*1*	y_7
8	+1	+1	+1	*+1*	*+1*	*+1*	*+1*	*1*	y_8

Comme pour un essai 2^2, on multiplie chacune de ces colonnes par la colonne des réponses – ligne par ligne – et on fait la somme.
La seule différence est qu'il faut diviser le résultat obtenu par **8** (au lieu de 4 avec un essai 2^2)

Exercice 3 : le calcul avec EXCEL conduit à :
$E_A = 4,41$ $E_B = 0,89$ $E_C = 3,89$ $E_{AB} = 1,86$ $E_{AC} = 0,36$ $E_{BC} = -0,81$ $E_{ABC} = 0,16$
$y_C = 7,94$

Plan 2^3 - matrice des effets – calcul matriciel

n°	A	B	C	AB	AC	BC	ABC	M	rép
1	-1	-1	-1	+1	+1	+1	-1	1	0
2	+1	-1	-1	-1	-1	+1	+1	1	4,7
	.	.	.	.	.	.	.	.	.
8	+1	+1	+1	+1	+1	+1	+1	1	18,7

matrice X

matrice Y

matrice E

8 équations avec 8 inconnues : E_A , E_B , E_C , E_{AB} , E_{AC} , E_{BC} , E_{ABC} , y_C

$$0 = -E_A - E_B - E_C + E_{AB} + E_{AC} + E_{BC} - E_{ABC} + y_C$$

$$4{,}7 = +E_A - E_B - E_C - E_{AB} - E_{AC} + E_{BC} + E_{ABC} + y_C$$

....

$$18{,}7 = +E_A + E_B + E_C + E_{AB} + E_{AC} + E_{BC} + E_{ABC} + y_C$$

$$y = E_A x_A + E_B x_B + E_C x_C + E_{AB} x_A x_B + E_{AC} x_A x_C + E_{BC} x_B x_C + E_{ABC} x_A x_B x_C + y_C$$

En écriture matricielle :

Y = X.E donc $\mathbf{E = X^{-1} . Y}$

1-3 : Le plan 2^n

1-3-1 : Conditions expérimentales – Éléments à calculer

Ils constituent la généralisation des plans 2^2 et 2^3.

- Les conditions expérimentales sont prises aux extrémités du domaine. Ces conditions deviennent rapidement nombreuses (expérimentation lourde…) car chaque fois que n augmente de 1 unité, le nombre de conditions est multiplié par 2.
- Le nombre d'éléments à calculer dépend de n, il y a :
  - $\overline{y}_C$ réponse calculée au centre du domaine. On notera qu'aucun essai n'est prévu dans ce type de plan pour cette condition.
  - n effets principaux
  - C_n^2 effets d'interaction du 2e ordre
  - C_n^3 effets d'interaction du 3e ordre
  - C_n^4 effets d'interaction du 4e ordre
  - etc…

Le plan 2^n (éléments à calculer)

- En pratique ces valeurs sont obtenues assez facilement avec la matrice des effets. Il convient simplement de numéroter convenablement les conditions (n° essai), d'utiliser les facteurs centrés réduits (-1, +1) et de construire la matrice progressivement.

- les **facteurs** :

  A : suite de –1 et +1 alternés , **B** : 2 (-1) puis 2 (+1) puis…
  C : 4 (-1) puis 4 (+1) puis… , **D** : 8 (-1) puis 8 (+1) puis…
  E : 16 (-1) puis 16 (+1) puis… , etc…
- les **colonnes d'interaction** sont obtenues en multipliant les colonnes des facteurs; par exemple l'interaction ABCD s'obtient en multipliant ligne par ligne A par B par C par D.
- la **colonne M** ne contient que des +1

- notons que l'ordre dans lequel sont établies les colonnes n'a pas d'incidence sur le calcul (par exemple on peut placer BC à gauche de AB ou l'inverse)
- en revanche les numéros de la colonne n° essai définissent la succession des conditions expérimentales et il est primordial de mettre la valeur mesurée correspondant à cette condition dans la colonne réponse.

1-3-2 : Plan 2^4

Exercice 4 : lors de l'optimisation des conditions expérimentales d'une réaction de précipitation, on a retenu comme facteurs à étudier susceptibles d'agir sur le poids de précipité pesé dans des conditions identiques :

facteurs	niveaux
A : température de la réaction	60°C – 70°C
B : concentration d'un réactif	1 $g.L^{-1}$ - 2 $g.L^{-1}$
C : temps de contact	30 min – 45 min
D : débit de lavage du précipité	1 $l.min^{-1}$ – 0,5 $l.min^{-1}$

- Soit 16 conditions expérimentales : il est décidé de ne pas faire de répétitions.
- les résultats des 16 expériences (poids de précipité) sont donnés ci-dessous :

Condition	Poids
$A_0B_0C_0D_0$	60,6
$A_1B_0C_0D_0$	61,0
$A_0B_1C_0D_0$	60,3
$A_1B_1C_0D_0$	61,7
$A_0B_0C_1D_0$	62,0
$A_1B_0C_1D_0$	61,5
$A_0B_1C_1D_0$	61,7
$A_1B_1C_1D_0$	62,4

Condition	Poids
$A_0B_0C_0D_1$	59,6
$A_1B_0C_0D_1$	61,1
$A_0B_1C_0D_1$	60,7
$A_1B_1C_0D_1$	61,3
$A_0B_0C_1D_1$	61,6
$A_1B_0C_1D_1$	61,9
$A_0B_1C_1D_1$	62,3
$A_1B_1C_1D_1$	62,8

Plan 2^4 – exemple 4 (suite)

Dans la matrice des effets il y a 16 colonnes :

- 4 effets principaux
- C_4^2 soit 6 interactions du 2e ordre
- C_4^3 soit 4 interactions du 3e ordre
- C_4^4 soit 1 interaction du 4e ordre
- 1 colonne M de +1

Exemple 4 (suite) – Calcul des effets

Le calcul des effets (ne pas oublier de diviser par 16 puisqu'il y a 16 essais) conduit à :

E_A=	0,306	E_{AB}=	0,093	E_{ABC}=	0,081	E_{ABCD}=	0,056
E_B=	0,244	E_{AC}=	-0,181	E_{ABD}=	-0,181		
E_C=	**0,619**	E_{AD}=	0,056	E_{ACD}=	0,019		
E_D=	0,006	E_{BC}=	0,031	E_{BCD}=	0,006		
		E_{BD}=	0,119				
		E_{CD}=	0,119			$\bar{y}_C$	61,406

• A l'examen de ces effets, on constate qu'ils sont faibles par rapport au poids moyen calculé 61,4 (maximum 0,6 pour le facteur C) et on peut se demander s'ils ne sont pas de l'ordre de grandeur de ceux qui se produisent par suite de la dispersion des valeurs expérimentales. Nous allons bientôt examiner cet aspect du problème.

• Un résultat est clair : l'effet de D (0,006) est pratiquement nul : le changement du mode de lavage ne joue pas sur le poids de précipité.

On peut donc réexaminer les résultats en ne tenant pas compte du facteur D : on est ramené à l'étude d'un plan 2^3 (A, B, C) chaque condition expérimentale ayant été répétée 2 fois.

1-3-3 : Calcul des effets d'un plan 2^n dans le cas où il y a des répétitions

Ce cas concerne :

- les plans factoriels pour lesquels au départ il a été prévu que chaque condition expérimentale serait répétée 2 fois, 3 fois, etc.…

Le but est d'améliorer la précision de la détermination des effets.

Cette pratique est très courante en BIOLOGIE parce que la dispersion des mesures est souvent grande. En revanche, elle est rare dans les essais industriels où la répétabilité des déterminations est généralement bonne.

- les plans factoriels pour lesquels après analyse, un facteur se révèle sans effet ainsi que nous venons de le voir.

Le calcul des effets utilise encore la matrice des effets.

Avec un tableur type EXCEL on peut opérer de 2 façons.

Calcul des effets d'un plan 2^n dans le cas où il y a des répétitions

1ère façon

On calcule les moyennes des réponses pour chaque condition expérimentale et on effectue le calcul classique des effets avec les moyennes.

Par exemple pour un plan 2^2 "tripliqué"

n° essai	A	B	AB	M
1	-1	-1	+1	+1
2	+1	-1	-1	+1
3	-1	+1	-1	+1
4	+1	+1	+1	+1

rép 1	rép 2	rép 3	moyenne
y_1	y_1'	y_1''	$\bar{y}_1$
y_2	y_2'	y_2''	$\bar{y}_2$
y_3	y_3'	y_3''	$\bar{y}_3$
y_4	y_4'	y_4''	$\bar{y}_4$

$$E_A = \frac{1}{4}(-\bar{y}_1 + \bar{y}_2 - \bar{y}_3 + \bar{y}_4) \qquad E_B = \frac{1}{4}(-\bar{y}_1 - \bar{y}_2 + \bar{y}_3 + \bar{y}_4)$$

$$E_{AB} = \frac{1}{4}(+\bar{y}_1 - \bar{y}_2 - \bar{y}_3 + \bar{y}_4)$$

Calcul des effets d'un plan 2^n dans le cas où il y a des répétitions

2ème façon

On peut aussi "copier" la matrice des effets autant de fois qu'il y a de répétitions et on effectue le calcul des effets à partir des réponses individuelles en divisant par le nombre d'expériences réalisées. Par exemple pour un plan 2^2 "tripliqué" :

n° essai	A	B	AB	M		réponse
1	-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

n° essai	A	B	AB	M		réponse
1	-1	-1	+1	+1	②	$y_1^{'}$
2	+1	-1	-1	+1		$y_2^{'}$
3	-1	+1	-1	+1		$y_3^{'}$
4	+1	+1	+1	+1		$y_4^{'}$

n° essai	A	B	AB	M		réponse
1	-1	-1	+1	+1	③	$y_1^{''}$
2	+1	-1	-1	+1		$y_2^{''}$
3	-1	+1	-1	+1		$y_3^{''}$
4	+1	+1	+1	+1		$y_4^{''}$

$$E_A = \frac{1}{12}(-y_1 + y_2 - y_3 + y_4 - y_1^{'} + y_2^{'} - y_3^{'} + y_4^{'} - y_1^{''} + y_2^{''} - y_3^{''} + y_4^{''})$$

idem pour E_B, E_{AB}, y_C

Calcul des effets d'un plan 2^n dans le cas où il y a des répétitions

2ème façon (suite)

Cette façon de calculer montre immédiatement que dans un essai 2^n comportant un facteur à effet nul et ramené à un plan 2^{n-1} comportant 2 répétitions, il n'y a pas besoin de refaire les calculs. Il suffit de supprimer dans la matrice des effets tout ce qui se rapporte au facteur à effet nul.

Dans l'exemple 4, on supprimera donc les colonnes D, AD, BD, CD, ABD, ACD, BCD et ABCD.

Les autres colonnes restent les mêmes et les effets sont donc identiques à ceux que l'on avait calculés, soit :

E_A=	0,306	E_{AB}=	0,093	E_{ABC}=	0,081
E_B=	0,244	E_{AC}=	-0,181	$\bar{y}_C$	61,406
E_C=	0,619	E_{BC}=	0,031		

Alors quel est l'intérêt de ne pas tenir compte de D pour avoir des répétitions ?

Il réside dans un autre aspect du problème, celui de la signification des effets : c'est ce que nous allons maintenant examiner.

2- ANALYSE DES RESULTATS : QUELS EFFETS (PRINCIPAUX, INTERACTIONS) SONT SIGNIFICATIFS ?

- Les exemples précédents ont montré qu'aucun effet calculé, même parmi les interactions d'ordres élevés, n'est rigoureusement égal à 0.
Cette constatation n'est pas surprenante puisque les effets sont calculés à partir des réponses mesurées, entachées d'erreur aléatoire.
Un effet, positif ou négatif, est d'autant plus crédible que sa valeur absolue est grande.

❑ A partir de quelle VALEUR SEUIL, peut-on dire qu'un effet est SIGNIFICATIF, autrement dit qu'il a certainement une existence réelle ?

- Soit σ_E l'écart-type(*) d'un effet principal ou d'interaction de grandeur E.

- si $|E|$ est très supérieur à σ_E , au moins 3 fois, l'existence de l'effet sera considéré comme certaine.

- si $|E|$ est très nettement inférieure à σ_E , l'effet calculé a plus de chances de résulter de la dispersion des mesures de réponses que de l'existence réelle de l'effet : la décision sera que l'effet n'existe pas.

- si $|E|$ et σ_E sont du même ordre de grandeur, c'est le cas critique pour lequel il faudra avoir recours aux tests statistiques pour décider de l'existence ou non de l'effet.

(*) pour des GRANDEURS CALCULEES (c'est le cas des effets) il faut préférer le terme d'ERREUR TYPE ou d'ERREUR STANDARD.

❑ Base théorique : relation entre variabilité des effets et variabilité des réponses individuelles.

• Dans un essai factoriel, n'importe quel effet $\mathbf{E_i}$ se calcule à partir des **n** réponses individuelles par $E_i = \frac{1}{n}(\pm y_1 \pm y_2 \pm \ldots\ldots\ldots \pm y_n)$ (+ ou – selon les effets)

• Comme les mesures y_1, y_2, ……, y_n sont indépendantes, les variances s'ajoutent

$$\sigma^2_{E_i} = \frac{1}{n^2}(\sigma^2_{y_1} + \sigma^2_{y_2} + \ldots\ldots + \sigma^2_{y_n})$$

$(\sigma^2_{y_1}, \sigma^2_{y_2}, \ldots, \sigma^2_{y_n}$ variances des mesures, $\sigma^2_{E_i}$ variance de l'effet $E_i)$

Dans la suite, nous supposerons que les réponses individuelles sont distribuées normalement et ont toutes le même écart-type σ_y dans le domaine expérimental : $\sigma^2_{y_1} = \sigma^2_{y_2} = \ldots = \sigma^2_{y_n}$

$$\sigma^2_{E_i} = \frac{1}{n^2}(\underbrace{\sigma^2_{y_n} + \sigma^2_{y_n} + \ldots\ldots + \sigma^2_{y_n}}_{\text{n fois}}) = \frac{\sigma^2_{y_n}}{n}$$

L'erreur type σ_E est la même pour tous les effets

$$\boxed{\sigma_E = \frac{\sigma_y}{\sqrt{n}}}$$

Erreur-type d'un effet

$$\boxed{\sigma_E = \frac{\sigma_y}{\sqrt{n}}}$$

- Cette formule est donc valable pour les effets principaux, les interactions et la réponse $\overline{y}_C$ au centre du domaine. La connaissance de σ_y permet donc d'estimer σ_E .

- Dans la pratique expérimentale, plusieurs cas peuvent se présenter.
  1. **l'écart-type σ_y est connu** par des expériences antérieures de même type que celles du plan et suffisamment nombreuses pour être fiables. C'est le meilleur des cas.
  2. **l'écart-type σ_y n'est pas connu** mais l'expérimentateur a prévu dans le plan quelques **expériences complémentaires** pour l'estimer. Un cas particulier est celui où le plan a été conçu pour comporter des répétitions.
  3. **l'écart-type σ_y n'est pas connu** et il n'y a eu ni répétitions du plan, ni essais complémentaires. On peut encore estimer σ_E , l'erreur-type des effets, en utilisant les interactions d'ordres élevés. C'est la méthode la moins satisfaisante.

2-1 : Détermination des effets significatifs quand σ_y connu par expériences antérieures

Reprenons l'exercice 3 (louche d'une solution)

facteurs	niveaux
A = θ re	20°C – 40°C
B = vitesse agitation	100 $t.min^{-1}$ - 300 $t.min^{-1}$
C = concentration additif	0,1 % - 0,5 %

L'expérimentateur avait constaté depuis longtemps la présence de ce louche lors des préparations successives dans les conditions habituelles (30°C ; 200 $t.min^{-1}$; 0,3 %). A chaque préparation, il mesurait l'opacité pour s'assurer qu'elle ne dépassait pas une limite fixée.

En étudiant la distribution de 50 mesures de louche dans ces conditions, il trouve que les valeurs semblent se répartir selon une loi Normale.

- de moyenne 7,85
- d'écart-type σ_y = 2,45

- écart-type sur 50 valeurs σ_y = 2,45 n = 2^3 = 8 réponses

> Erreur-type d'un effet $$\boxed{\sigma_E = \frac{\sigma_y}{\sqrt{n}} = \frac{2,45}{\sqrt{8}} = 0,866 \approx 0,9}$$

• Pour savoir si la dispersion des mesures peut expliquer ou non la différence E – 0 d'un effet on calcule :

$$z_0 = \frac{|E|}{\sigma_E}$$

que l'on compare à la valeur seuil de **z** sur la table de la loi Normale réduite, le risque étant choisi.

• si $z_0 \geq 2$ on prend un risque de 5 % en affirmant que l'effet existe
• si $z_0 \geq 2,6$ -------------------------- 1 % ------------------------------------

On peut aussi exprimer le résultat du test en degré de signification

Exemple 3 (étude du louche)

$E_A = 4{,}41$ $E_B = 0{,}89$ $E_C = 3{,}89$ $E_{AB} = 1{,}86$ $E_{AC} = 0{,}36$ $E_{BC} = -0{,}81$
$E_{ABC} = 0{,}16$ $y_C = 7{,}94$

• Dans cette étude, seuls 3 effets peuvent être considérés comme significatifs au risque 5 % : **$E_A = 4{,}41$ $E_C = 3{,}89$ $E_{AB} = 1{,}86$**

- effet principal de la température (A) : $$z_0 = \frac{4,41}{0,866} = 5,09$$

- effet principal de l'additif (C) : $$z_0 = \frac{3,89}{0,866} = 4,49$$

- effet de l'interaction Θ^{re} X agitation (AB) : $$z_0 = \frac{1,86}{0,866} = 2,15$$

CONCLUSION

• L'augmentation de la concentration d'additif augmente l'opacité.
• L'augmentation de la température provoque aussi une augmentation du louche mais celle ci dépend de la vitesse d'agitation.

Exemple 3 PRESENTATION DES RESULTATS

A (Θ^{re})	B (agit)	réponse
-1	-1	$y_1=0$
+1	-1	$y_2=4{,}7$
-1	+1	$y_3=0$
+1	+1	$y_4=11{,}5$
-1	-1	$y_5=9$
+1	-1	$y_6=14{,}5$
-1	+1	$y_7=5{,}1$
+1	+1	$y_8=18{,}7$

Effet température :

- effet moyen

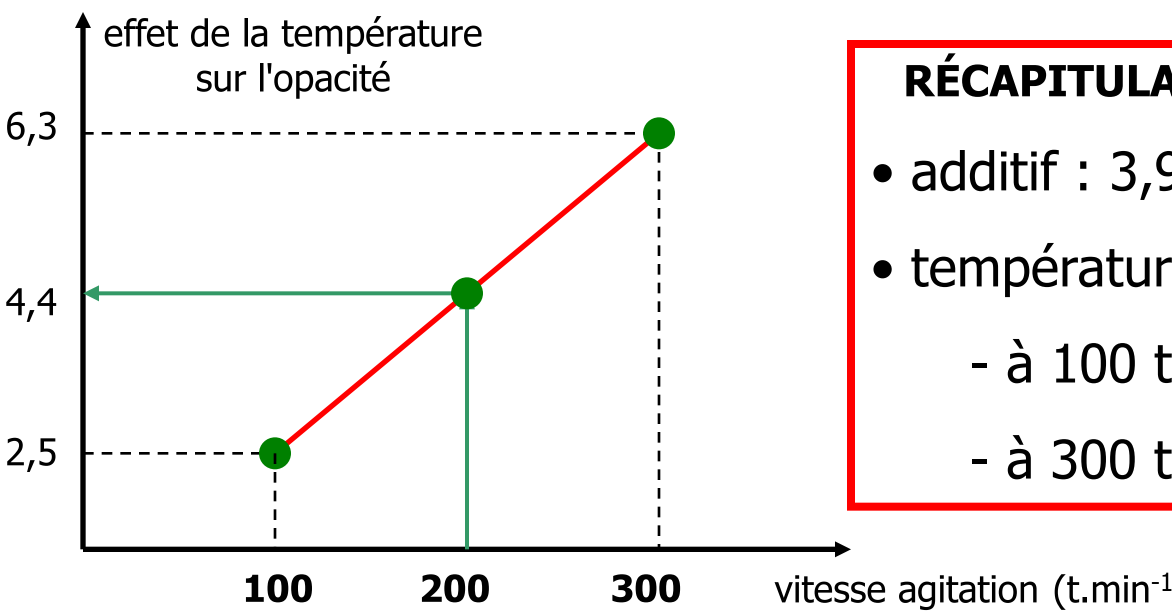
$$E_A = \frac{1}{8}(-y_1 + y_2 - y_3 + y_4 - y_5 + y_6 - y_7 + y_8) = 4{,}41 \qquad \sigma_E = \frac{\sigma_y}{\sqrt{n}} = \frac{2{,}45}{\sqrt{8}} = 0{,}9$$

- effet niveau BAS de B (100 t.min^{-1})

$$E_A = \frac{1}{4}(-y_1 + y_2 - y_5 + y_6) = 2{,}55 \approx 2{,}5 \qquad \sigma_E = \frac{\sigma_y}{\sqrt{n}} = \frac{2{,}45}{\sqrt{4}} = 1{,}23 \approx 1{,}2$$

- effet niveau HAUT de B (300 t.min^{-1})

$$E_A = \frac{1}{4}(-y_3 + y_4 - y_7 + y_8) = 6{,}275 \approx 6{,}3 \qquad \sigma_E = \frac{\sigma_y}{\sqrt{n}} = \frac{2{,}45}{\sqrt{4}} = 1{,}23 \approx 1{,}2$$



RÉCAPITULATIF DES VALEURS D'EFFET

- additif : 3,9 ± 0,9 unités d'opacité
- température :
  - à 100 t.min^{-1} : 2,5 ± 1,2 unités
  - à 300 t.min^{-1} : 6,3 ± 1,2 unités

REMARQUE

❑ Bien que la variabilité des réponses ne soit pas évaluée en même temps que l'expérience factorielle, cette technique est souvent utilisée dans le domaine de la **Chimie et des essais industriels** et elle donne des résultats satisfaisants.

Cela tient au fait que, dans ces secteurs d'expérimentation, il y a souvent une bonne stabilité de la dispersion dans le temps.

Elle présente en revanche des risques d'erreur d'analyse dans les sciences du vivant (**Biologie**) à cause de l'hétérogénéité du matériel expérimental et de son inconstance dans le temps ; il est meilleur, en général, d'estimer la variabilité en même temps que l'expérience factorielle.

Rappel : pour estimer l'erreur-type d'un effet

$$\sigma_E = \frac{\sigma_y}{\sqrt{n}}$$

- Dans la pratique expérimentale, plusieurs cas peuvent se présenter.
  1. **l'écart-type σ_y est connu** par des expériences antérieures de même type que celles du plan et suffisamment nombreuses pour être fiables. C'est le meilleur des cas.
  2. **l'écart-type σ_y n'est pas connu** mais l'expérimentateur a prévu dans le plan quelques **expériences complémentaires** pour l'estimer. Un cas particulier est celui où le plan a été conçu pour comporter des répétitions.
  3. **l'écart-type σ_y n'est pas connu** et il n'y a eu ni répétitions du plan, ni essais complémentaires. On peut encore estimer σ_E , l'erreur-type des effets, en utilisant les interactions d'ordres élevés. C'est la méthode la moins satisfaisante.

2-2 : Détermination des effets significatifs au moyen d'expériences complémentaires

2-2-1 : Quelques essais complémentaires

On prévoit n répétitions (n essais, au minimum 5, en un même point, de préférence au centre du domaine) qui permettent d'estimer l'écart type s_y d'une réponse y puis de calculer s_E , l'erreur type d'un effet, grâce à la relation théorique précédente.

Chaque effet (principal et interaction) est alors testé par un **test t de Student** à n-1 degrés de liberté :

$$t_C = \frac{E}{s_E} \quad \text{comparé à} \quad t_{Seuil} \text{ (table de student au risque 0,05)}$$

Lorsque **le degré de signification du test p est < à 0,05** , on considère l'effet comme ayant une existence réelle : **le facteur agit sur la réponse.**

2-2-1 : Quelques essais complémentaires (suite)

Exercice 2 (question 2)

Au centre du domaine (70°C, 12,5 g.L-1), il a été effectué 6 essais complémentaires, on a obtenu les rendements (en %) : 77,3 – 79,1 – 77,8 – 77,0 – 77,7 – 79,1

- écart-type estimé à 5 d.d.l. $s_y = 0,899$ $n = 2^2 = 4$ réponses

> Erreur-type d'un effet

$$\boxed{s_E = \frac{s_y}{\sqrt{n}} = \frac{0,899}{\sqrt{4}} = 0,449}$$

• Pour savoir si la dispersion des mesures peut expliquer ou non la différence E – 0 d'un effet on calcule : $t_0 = \frac{|E|}{s_E}$

que l'on compare à la valeur seuil de **t** sur la table de STUDENT, le risque étant choisi.

• si $t_0 \geq t$ on prend un risque de 5 % en affirmant que l'effet existe

On peut aussi exprimer le résultat du test en degré de signification

$$t_{seuil} = 2,57 \text{ (5 d.d.l)}$$

• Dans cette étude, les 3 effets peuvent être considérés comme significatifs au risque 5 %

- effet principal de la température (A) : *très significatif*

$$t_0 = \frac{6,25}{0,45} = 13,9$$

- effet principal de la concentration (B) : *très significatif*

$$t_0 = \frac{11,25}{0,45} = 25,0$$

- effet de l'interaction θ re X conc (AB) : *significatif*

$$t_0 = \frac{1,25}{0,45} = 2,78$$

RÉCAPITULATIF DES VALEURS D'EFFET

- température : 11,25 ± 0,45
- concentration : 6,25 ± 0,45
- interaction θ re X conc : 1,25 ± 0,45

2-2 : Détermination des effets significatifs au moyen d'expériences complémentaires

2-2-2 : Répétition du plan factoriel

On a n réponses pour chaque condition expérimentale.

On peut calculer :
- la dispersion (variance) dans chaque groupe
- la variance s^2_y commune

qui permet d'estimer l'écart type s_y d'une réponse y puis de calculer s_E ,

Chaque effet (principal et interaction) est alors testé comme précédemment par un **test t** :

2-2-2 : Répétition du plan factoriel (suite)

Exercice 4 (question 3)

Nous avons vu (question 2) qu'on pouvait considérer qu'il s'agit d'une expérience à 3 facteurs (température, concentration, temps), **chaque condition étant répétée 2 fois**.

On estime la variance sur chacun des 8 groupes (8 conditions), chacune à **1 d.d.l.** et l'on en déduit l'estimation de la variance commune :

$$s^2_y = 0,1356 \text{ à 8 d.d.l.}$$

- écart-type estimé à 8 d.d.l. $s_y = 0,368$ N $= 2 \times 2^3 = 16$ réponses

> Erreur-type d'un effet

$$\boxed{s_E = \frac{s_y}{\sqrt{n}} = \frac{0,368}{\sqrt{16}} = 0,092}$$

t_{seuil} **= 2,31** (8 d.d.l)

• Dans cet exemple, seuls les 3 effets principaux sont significatifs au risque 5 %

- effet principal de la température (A) : *significatif* $$t_0 = \frac{0,306}{0,092} = 3,33$$

- effet principal de la concentration (B) : *significatif* $$t_0 = \frac{0,244}{0,092} = 2,65$$

- effet principal du temps de contact (C) : *très significatif* $$t_0 = \frac{0,619}{0,092} = 6,72$$

RÉCAPITULATIF DES VALEURS D'EFFET

- l'augmentation de 5°C de la température augmente le rendement de : 0,31 ± 0,09
- l'augmentation de 0,5 g/L de la concentration augmente le rendement de : 0,24 ± 0,09
- l'augmentation de 7,5 min du temps de contact augmente le rendement de : 0,62 ± 0,09

Ces facteurs semblent agir indépendamment

REMARQUE : Il est aussi possible pour tester les effets d'utiliser l'ANALYSE DE VARIANCE (à 3 facteurs croisés ici). *On obtient bien sur les mêmes degrés de signification. Calcul avec SigmaStat :*

SigmaStat - [Data 1]

File Edit View Format Tools Graph Statistics Transforms Window Help

Three Way ANOVA

	1-A	2-B	3-C	4-reponse	5
1	-1.0000	-1.0000	-1.0000	60.6000	
2	1.0000	-1.0000	-1.0000	61.0000	
3	-1.0000	1.0000	-1.0000	60.3000	
4	1.0000	1.0000	-1.0000	61.7000	
5	-1.0000	-1.0000	1.0000	62.0000	
6	1.0000	-1.0000	1.0000	61.5000	
7	-1.0000	1.0000	1.0000	61.7000	
8	1.0000	1.0000	1.0000	62.4000	
9	-1.0000	-1.0000	-1.0000	59.6000	
10	1.0000	-1.0000	-1.0000	61.1000	
11	-1.0000	1.0000	-1.0000	60.7000	
12	1.0000	1.0000	-1.0000	61.3000	
13	-1.0000	-1.0000	1.0000	61.6000	
14	1.0000	-1.0000	1.0000	61.9000	
15	-1.0000	1.0000	1.0000	62.3000	
16	1.0000	1.0000	1.0000	62.8000	
17					

Three Way Analysis of Variance mercred

Data source: Data 1 in ex4-3.SNB

Balanced Design

Dependent Variable: reponse

Normality Test: Passed (P > 0.050)

Equal Variance Test: Failed (P = <0.001)

Source of Variation	DF	SS	MS	F	P
A	1	1.501	1.501	11.065	0.010
B	1	0.951	0.951	7.009	0.029
C	1	6.126	6.126	45.166	<0.001
A x B	1	0.141	0.141	1.037	0.338
A x C	1	0.526	0.526	3.876	0.085
B x C	1	0.0156	0.0156	0.115	0.743
A x B x C	1	0.106	0.106	0.779	0.403
Residual	8	1.085	0.136		
Total	15	10.449	0.697		

Rappel : pour estimer l'erreur-type d'un effet

$$\boxed{\sigma_E = \frac{\sigma_y}{\sqrt{n}}}$$

- Dans la pratique expérimentale, plusieurs cas peuvent se présenter.

1. **l'écart-type σ_y est connu** par des expériences antérieures de même type que celles du plan et suffisamment nombreuses pour être fiables. C'est le meilleur des cas.
2. **l'écart-type σ_y n'est pas connu** mais l'expérimentateur a prévu dans le plan quelques **expériences complémentaires** pour l'estimer. Un cas particulier est celui où le plan a été conçu pour comporter des répétitions.
3. **l'écart-type σ_y n'est pas connu** et il n'y a eu ni répétitions du plan, ni essais complémentaires. On peut encore estimer σ_E , l'erreur-type des effets, en utilisant les interactions d'ordres élevés. C'est la méthode la moins satisfaisante.

2-3 : Détermination des effets significatifs sans essais complémentaires

2-3-1 : à l'aide des interactions d'ordres élevés

Il faut supposer que les interactions d'ordre élevé (3 et plus) n'ont pas d'effet et que les valeurs calculées ne diffèrent de 0 que par suite de la dispersion expérimentale des mesures de réponses. Ces valeurs constituent donc des mesures d'écarts à 0 aléatoires et permettent donc d'estimer l'erreur type s_E , valable pour tous les effets.

On teste la signification des effets intéressants (principaux et de 2ème ordre) par une série de tests t de Student , comme précédemment.

La méthode n'est fiable que s'il y a suffisamment d'interactions d'ordres supérieurs dans l'estimation de s_E. **Le plan doit donc comporter au moins 4 facteurs.**

Exemple : exercice n°4 , question 4.

exercice n°4 , question 4.

- Effets dont on teste la signification (jusqu'à interaction ordre 2) :

A, B, C, D, AB, AC,.......................CD

- Effets supposés nuls* et servant à calculer la dispersion (interaction ordre 3 et 4) :

ABC, ABD,.......................ABCD

*_remarque : *moyenne calculée 0,004*

estimation de σ^2_E = moyenne des carrés des écarts à la moyenne 0

$$\boxed{s_E^2 = \frac{\sum_1^5 (E_i - 0)^2}{5} = \frac{E^2_{ABC} + E^2_{ABD} + \ldots + E^2_{ABCD}}{5} = 0{,}086}$$

$$\boxed{S_E = 0{,}0927 \text{ (5 d.d.l.)}}$$

<u>REMARQUE</u> : S_E a pratiquement la même valeur que celle obtenue en négligeant D (question 3) mais il n'y a ici que 5 d.d.l. au lieu de 8, les degrés de signification sont plus élevés et certaines conclusions pourraient « basculer »
Ainsi, on aboutit aux mêmes conclusions (A, B et C significatifs) mais de justesse pour B.

2-3 : Détermination des effets significatifs sans essais complémentaires

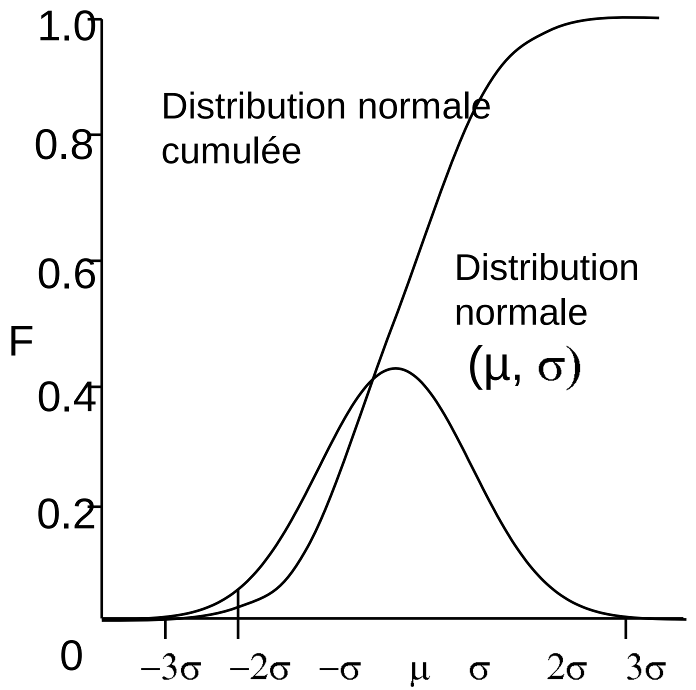
2-3-2 : critère graphique de normalité

Les effets inactifs sont supposés être distribués normalement avec comme moyenne 0 et comme écart type σ_E. Ils doivent donc conduire à des points sensiblement alignés dans un graphe de probabilité normale.

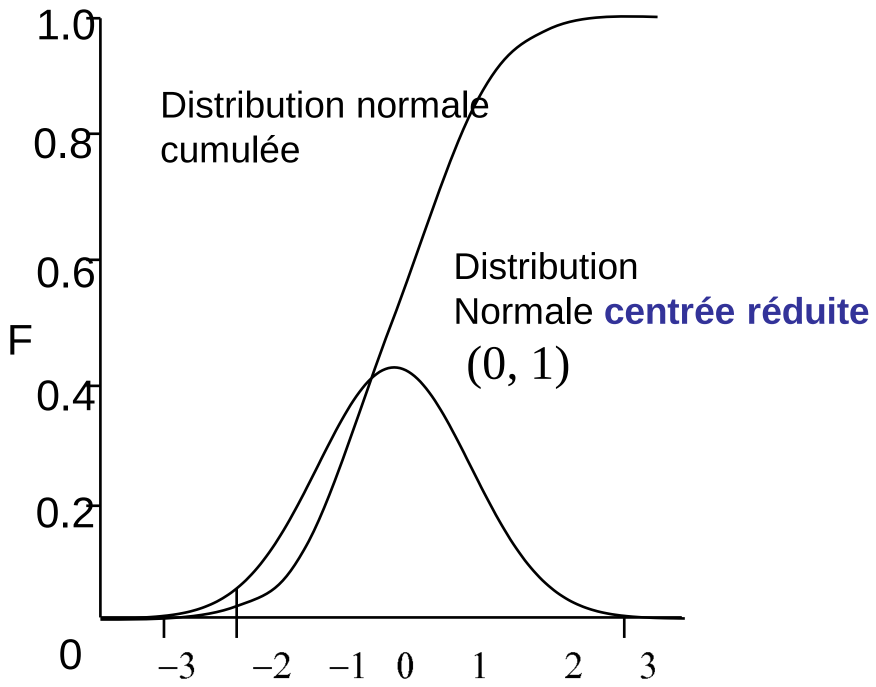
Les effets significatifs (non nuls) seront situés en dehors de cette droite. Ils se comporteront comme des valeurs aberrantes

Exemple : exercice n°5 , question 4.

Critère graphique de normalité (principe)

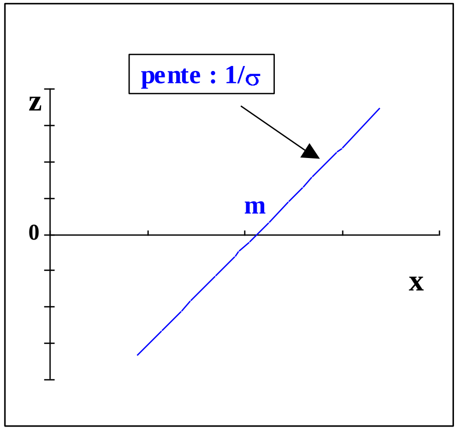


X suit une loi Normale (μ, σ)



Z suit une loi Normale (0, 1)

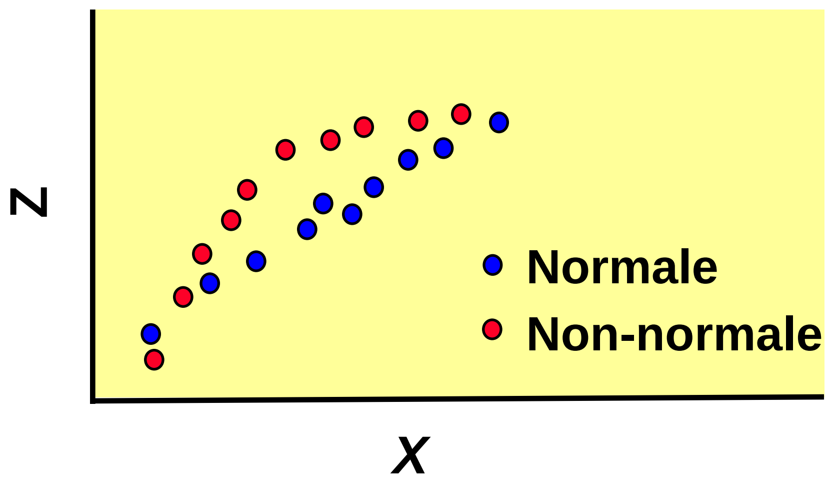
Relation linéaire : $$Z_i = \frac{X_i - \mu}{\sigma}$$



Critère graphique de normalité (principe)

Pour tester la normalité d'une série de valeurs X on calcule pour chacune l'équivalent de la probabilité relative cumulée :
- La Fréquence Relative Cumulée (FRC)

on lui fait correspondre par la table de la loi Normale la valeur Z correspondante :
On représente Z=f(X), si les points sont sensiblement alignés, l'hypothèse de normalité est acceptable



Critère graphique de normalité (calculs)

Ranger les données xi par valeurs croissantes ; noter 1 pour la plus petite, 2 pour la deuxième, ..., et N pour la plus grande. Ces nombres sont appelés les rangs (1, 2, ..., ri, ...N).

On admettra la formule suivante :

$$\mathrm{FRC}_i = \frac{r_i - 3/8}{N + 1/4}$$

La table de la loi Normale permet d'obtenir la valeur z_i correspondant à FRCi (voir exemples dans fichier Excel)

application aux effets : exercice n°5 , question 4.

Méthodologie dans l'étude d 'un phénomène

Lors de l 'étude d 'un phénomène, plusieurs questions se posent, auxquelles répondent différents types de plans. On peut distinguer 3 grandes étapes dans l 'acquisition des connaissances :

➊ Recherche des facteurs influents traité !

➋ Modélisation

➌ Optimisation

⇒ la modélisation

Quand les facteurs influents ont été identifiés et leur importance quantifiée, on recherche ensuite l'équation permettant de décrire les variations de la réponse étudiée en fonction de celles des facteurs influents ; cette seconde étape constitue la ***modélisation***.

Modèles utilisés : linéaires polynomiaux du 1er degré ou du 2e degré :

$$\hat{y} = y_C + a_1X_1 + a_2X_2 + a_{12}X_1X_2....$$

$$\hat{y} = y_C + a_1X_1 + a_2X_2 + a_{12}X_1X_2 + a_{11}x^2{}_1 + a_{22}x^2{}_2.....$$

Nous nous limiterons aux polynômes du 1er degré

3 - LE MODELE LINEAIRE (1er degré)

La méthode précédente de calcul des effets utilise de façon sous-jacente un *modèle linéaire :* *u*ne équation permet de représenter la réponse Y en fonction des facteurs X_A , X_B …

Ce modèle permet de quantifier les effets E_A , E_B , E_{AB} .. et de détecter ceux qui sont significatifs.

Cette équation a d'autres objectifs :

- permettre de prévoir la réponse dans des conditions expérimentales où aucune mesure n'a été effectuée (à l'intérieur du domaine).
- servir de point de départ dans une étude d'optimisation.

3-1 : l’équation du modèle linéaire (plans 2^n)

Dans le modèle linéaire , la réponse prédite, notée $\hat{y}$ est la somme de plusieurs termes :

- la moyenne de l’ensemble des réponses du plan $\bar{y}_c$
- pour un facteur, le terme est une fonction du 1er degré par rapport au facteur centré réduit **X** , le coefficient de proportionnalité étant égal à l’effet principal de ce facteur (effet moyen).
- pour une interaction , le terme est constitué par les produits des facteurs centrés réduits intervenant dans l’interaction et l’effet moyen de cette interaction.

$$\hat{y} = \bar{y} + E_A X_A + E_B X_B + E_{AB} X_A X_B + \ldots$$

Cas d'un plan 2^2 avec interaction

Exemple : exercice n°2 pour lequel E_A , E_B et E_{AB} sont significatifs :

$$\hat{y} = \bar{y} + E_A X_A + E_B X_B + E_{AB} X_A X_B$$

Les effets calculés ont pour valeurs

$E_A = 6,25$ $E_B = 11,25$ $E_{AB} = 1,25$ avec $\bar{y} = 76,25$

Numériquement , pour des rendements y exprimés en % , on écrit :

$$\hat{Y} = 76,25 + 6,25\, X_A + 11,25\, X_B + 1,25\, X_A X_B$$

Cette équation permet de prédire les valeurs des rendements pour des conditions expérimentales situées à l'intérieur du domaine exploré par le plan.

Prédiction d'une réponse

Exemple : exercice n°2

$\hat{Y} = 76{,}25 + 6{,}25\, X_A + 11{,}25\, X_B + 1{,}25\, X_A\, X_B$

Quelle réponse peut-on prédire pour la condition expérimentale suivante : $\theta_A = 76°C$, $C_B = 11{,}5 g.L^{-1}$?

A la température θ_A *= 76°C correspond la valeur du facteur centré réduit (60<* θ<80):

$$X_A = \frac{76-70}{10} = +0{,}6$$

A la concentration C_B *= 11,5g.L*$^{-1}$ *correspond la valeur du facteur centré réduit (10<* C<15):

$$X_B = \frac{11{,}5-12{,}5}{2{,}5} = -0{,}4$$

La réponse prédite s'écrit :

$\hat{Y} = 76{,}25 + (6{,}25*0{,}6) + (11{,}25*-0{,}4) + (1{,}25*0{,}6*-0{,}4)$

$\hat{Y}$ = 75,20 %

Généralisation : cas d'un plan 2^n

Même pour un nombre n de facteurs élevés, le nombre d'effets ***significatifs*** *se limite souvent à* ***quelques facteurs et interactions de 2e ordre*** *: l'équation du modèle reste relativement simple*

Exemple : exercice n°5 (question 5)   Plan 2^5

Nous avons mis en évidence 3 facteurs significatifs : A, D, F

et 2 interactions significatives : AD, DF

$E_A = 10{,}25$	$E_D = 5{,}125$	$E_F = -3{,}5$
$E_{AD} = 6{,}125$	$E_{DF} = -5{,}125$	$\bar{y}_c = 116$

d'où l'équation :

$$\hat{y} = \bar{y} + E_A X_A + E_D X_D + E_F X_F + E_{AD} X_A X_D + E_{DF} X_D X_F$$

soit :

$$\boxed{\hat{y} = 116 + 10{,}25 * X_A + 5{,}125 * X_D - 3{,}5 * X_F + 6{,}125 * X_A X_D - 5{,}125 * X_D X_F}$$

Prédiction exercice n°5 (question 5, suite)

A -> 70 -> valeur du facteur centré réduit: $X_A = \frac{70-60}{20} = +0,5$

D -> 20 -> valeur du facteur centré réduit: $X_D = \frac{20-15}{5} = +1$

F -> 0,40 -> valeur du facteur centré réduit: $X_F = \frac{0,40-0,40}{0,20} = 0$

$$\boxed{\hat{y} = 116 + 10,25 * X_A + 5,125 * X_D - 3,5 * X_F + 6,125 * X_A X_D - 5,125 * X_D X_F}$$

soit : $\hat{y} = 129.3$

3-2 : Utilisation de la régression linéaire multiple

La régression linéaire multiple a pour objet d'expliquer une variable (ici Y) par plusieurs variables explicatives (ici les facteurs et les interactions) au moyen d'une relation linéaire :

$$\hat{y} = a + b_1 X_1 + b_2 X_2 + \ldots + b_n X_n$$

Elle convient donc parfaitement bien à la modélisation des plans factoriels 2^n comportant des facteurs quantitatifs :

- les variables explicatives sont indépendantes entre elles (par construction du plan)

- les coefficients de l'équation (effets estimés) également

Régression linéaire multiple (en pratique)

En pratique, dans la matrice des effets, on choisit comme variables explicatives les colonnes correspondant aux **facteurs** et aux **interactions d'ordre 2**.

On suppose donc que les interactions d'ordre élevé (3 et plus) n'ont pas d'effet (Cf **2-3-1**).

Rappelons que l'on peut ainsi estimer l'erreur type s_E , sur les effets et tester les effets principaux et les interactions d'ordre 2

NB : Les résultats peuvent en plus comprendre l'étude des ***résidus****, utiles pour tester la validité du modèle choisi (voir plus loin)*

Régression linéaire multiple (exemple)

<u>Exemple</u> : exercice n°4 (question 5)

L'on peut utiliser l'<u>utilitaire d'analyse</u> d'**Excel** (module régression linéaire).
On limite le calcul aux 4 facteurs A, B, C, D et aux 6 interactions d'ordre 2.

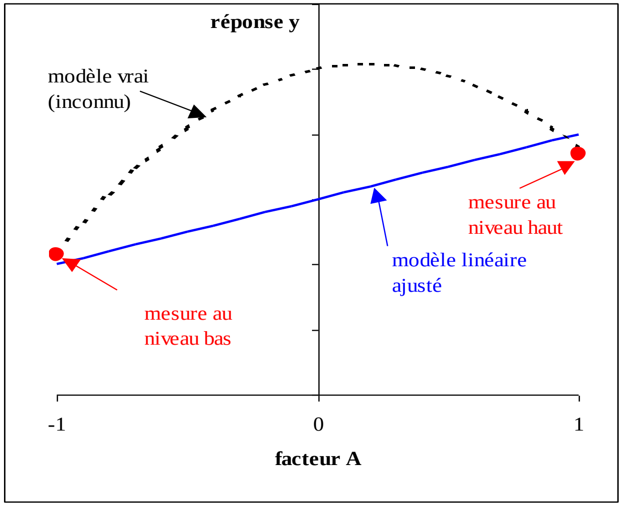
On retrouve les valeurs des effets (*coefficients*), de l'erreur type et le degré de signification (*probabilité*) des effets, déjà obtenus précédemment (beaucoup plus laborieusement !)

3-3 : validité de l'équation de prédiction

3-3-1 : test du modèle

- Dans les plans factoriels 2^n chaque facteur n'est expérimenté qu'en 2 points , aux extrémités –1 et +1 du domaine.

Cela explique le choix empirique d'une équation du 1er degré en fonction des facteurs centrés réduits X : par 2 points , il passe une droite et une seule ; mais il faut aussi ajouter qu'il peut y passer une infinité de courbes d'équations diverses …



- Il est donc important de s'assurer de la linéarité de l'équation dans tout le domaine en réalisant des expériences complémentaires avec un 3ème point à l'intérieur du domaine. On choisit généralement le centre (X = 0) situé à égale distance des extrémités expérimentées.

Quand on dispose de **répétitions au centre** du domaine expérimental, il est possible de **juger la linéarité** de l'équation de prédiction en comparant statistiquement

- la moyenne y_0 de ces répétitions.
- la réponse prédite au centre par l'équation linéaire , égale à la moyenne $\bar{y}$ des n réponses du plan factoriel 2^n.

Ces 2 moyennes doivent être en théorie égales et en pratique peu différer lorsque le modèle linéaire est valide.
Et, Lorsqu'on obtient des valeurs très différentes , s'écartant de plusieurs fois l'écart type , cela signifie que le modèle linéaire n'est pas valable et qu'il faut envisager un modèle empirique plus complexe , où les facteurs centrés réduits interviennent au 2ème degré par exemple.

Notons qu'il existe des tests statistiques d'écart à la linéarité du modèle

Exemple : exercice n°2 , 4ème question.

Exemple : exercice n°2 , 4ème question.

Erreur-type : 0,45

moyenne calculée : $\mathbf{y_c = 76,25 \pm 0,45}$ ($m \pm S_E$)

6 essais au centre ont été effectués pour lesquels la réponse moyenne des mesures est

moyenne mesurée : $\mathbf{y_0 = 78 \pm 0,37}$

Il y a désaccord entre la vraie moyenne mesurée et son évaluation par calcul. ; un test t permet de le confirmer :

d = 78 – 76.25 = 1.75

$s_d^2 = (0,37)^2 + (0,45)^2$

$S_d = 0,58$ (5+5=10 ddl)

t = d/ Sd = 3,02 (seuil = 2,23)

Le modèle du 1er degré n'est pas valide ici : toutes les prévisions de réponses sont fausses : il aurait fallu valider le modèle avant !

3-3 : validité de l'équation de prédiction

3-3-2 : étude des résidus

résidu = réponse mesurée – réponse prédite

$$r_i = y_i - \hat{y}_i \quad \textit{pour l'essai n° i}$$

Exemple : exercice n°5

$$\hat{y} = 116 + 10{,}25 * X_A + 5{,}125 * X_D - 3{,}\ 5 * X_F + 6{,}125 * X_A X_D - 5{,}125 * X_D X_F$$

n° essai	mesuré	prédit	résidu
1	109	105,125	*3,875*
2	113	113,375	*-0,375*
3	103	105,125	*-2,125*
4	113	113,375	*-0,375*
5	103	105,125	*-2,125*
...	...	...	...

- La notion de **résidu** n'a pas de sens si l'équation du modèle tient compte de tous les effets calculés : les résidus sont évidemment tous nuls (pour un plan sans répétition).

- Si l'on ne tient compte que des effets significatifs, les résidus ont des valeurs non nulles qui doivent être considérées comme des « termes d'erreur »

  résidu ←----→ partie de la mesure non explicable par le modèle

causes possibles :

- variations des facteurs non contrôlés pendant l'expérience
- imprécision de la méthode de mesure
- modèle inadapté
- etc…

- Par construction les résidus ont toujours comme **moyenne 0** et comme **écart-type S_y**, l'écart-type des réponses individuelles

- Les suppositions **statistiques** du modèle nécessitent que les résidus soient **distribués de façon aléatoire, suivant une loi normale**.

Méthodes de vérification de la Normalité et du caractère aléatoire :

- graphe de probabilité normale

- l'étude des résidus réduits (rri = ri / Sy doit être compris entre -2 et +2)

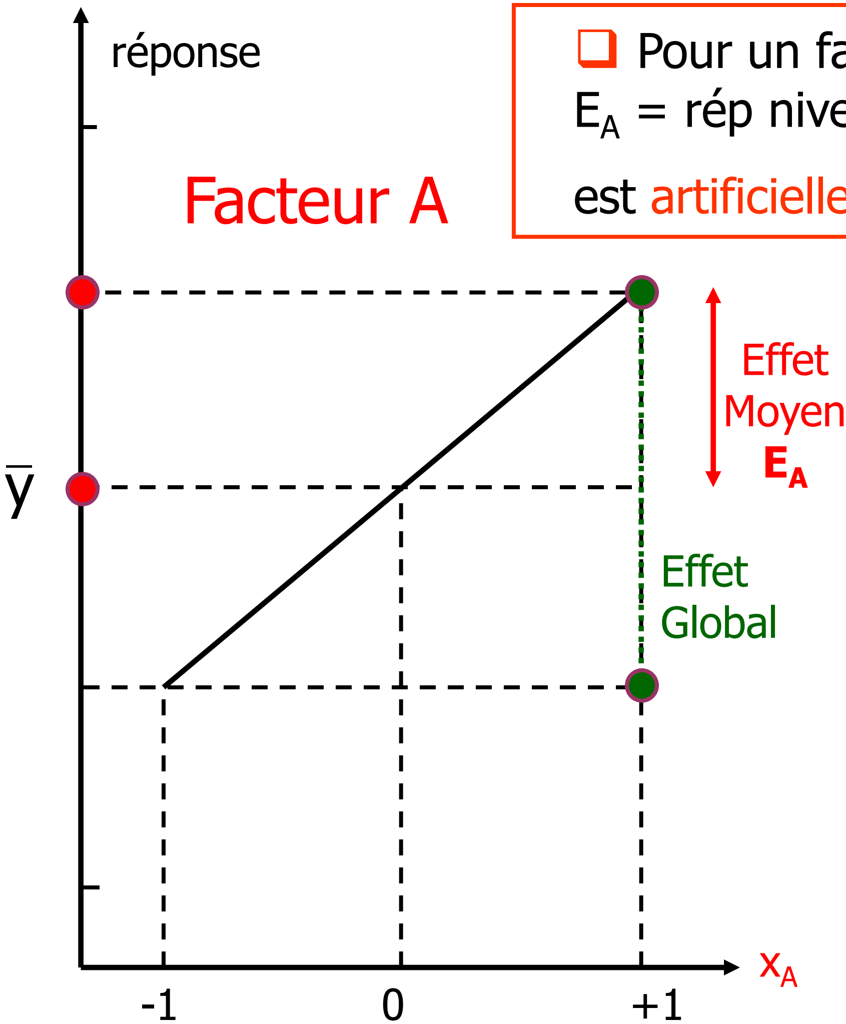
- etc…

Exemple : exercice n°5 , 6ème question.

3-4 : Modélisation en représentation qualitative

3-4-1 : Ecriture du modèle

- Pour un facteur qualitatif*, la définition de l'effet moyen E_A = rép niveau haut – moy générale des réponses $\bar{y}$ est artificielle (* exemple : présence ou non d'un catalyseur)



Seule la variation de réponse entre le niveau bas et le niveau haut présente de l'intérêt : c'est l'effet global 2 E_A
On peut donc distinguer un <u>effet moyen pour chaque niveau</u> de facteur :

- niveau haut : $+E_A$ = rép niveau haut – $\bar{y}$
- niveau bas : $-E_A$ = rép niveau bas – $\bar{y}$

Dans le modèle ce facteur sera représenté par : $[-E_A \;\; +E_A]\,[A]$

Facteur : $[-E_A \quad +E_A]\,[A]$

effet niveau bas - effet niveau haut

ce symbole est seulement présent pour indiquer que les valeurs d'effets se réfèrent à ce facteur

exemple : modèle avec 2 facteurs A et B significatifs sans interaction

$$\hat{y} = \bar{y} + [-E_A \quad +E_A]\,[A] + [-E_B \quad +E_B]\,[B]$$

l'avantage d'un tel modèle est que tous les effets des facteurs significatifs à chaque niveau sont représentés numériquement : on peut donc calculer facilement la réponse prédite pour un état donné

Interaction : pour une interaction de 2e ordre AB il y a 4 états possibles et on écrira le terme correspondant sous forme matricielle :

$$\begin{array}{l c}
\underline{\text{facteur } A} & \\
\text{niveau bas} & \multirow{2}{*}{$\begin{bmatrix} +E_{AB} & -E_{AB} \\ -E_{AB} & +E_{AB} \end{bmatrix} AB$} \\
\text{niveau haut} & \\
 & \begin{array}{cc} \text{niveau bas} & \text{niveau haut} \end{array} \\
 & \underline{\text{facteur } B}
\end{array}$$

exemple : modèle complet 2^2 avec interaction

$$\hat{y} = \bar{y} + [-E_A \quad +E_A][A] + [-E_B \quad +E_B][B] + \begin{bmatrix} +E_{AB} & -E_{AB} \\ -E_{AB} & +E_{AB} \end{bmatrix} AB$$

exemple numérique : $\bar{y} = 58{,}9$ $E_A = -25{,}0$ $E_B = +13{,}6$ $E_{AB} = +8{,}2$

Prédire la réponse pour A au niveau haut et B au niveau bas

Le modèle s'écrit :

$$\hat{y} = 58{,}9 + [+25 \quad -25][A] + [-13{,}6 \quad +13{,}6][B] + \begin{bmatrix} +8{,}2 & -8{,}2 \\ -8{,}2 & +8{,}2 \end{bmatrix} AB$$

Réponse prédite :

$$\hat{y} = 58{,}9 - 25 - 13{,}6 - 8{,}2 = \mathbf{12{,}1}$$

-Ce type de modélisation est généralisable à plus de 2 facteurs qualitatifs. On n'y fait figurer que les effets et les interactions d'ordre 2 significatifs

- Il peut s'appliquer aussi aux facteurs quantitatifs quand on ne s'intéresse qu'aux réponses prédites aux points expérimentés.

Exemple : exercice n°5 , 7ème question a)-

Plan 2^5 : nous avons mis en évidence 3 facteurs significatifs : A, D, F et 2 interactions significatives : AD, DF

$E_A = 10{,}25$	$E_D = 5{,}125$	$E_F = -3{,}5$
$E_{AD} = 6{,}125$	$E_{DF} = -5{,}125$	$\bar{y}_c = 116$

$$\hat{y} = 116 + [-10{,}25 \ \ +10{,}25\,][A\,] + [-\,5{,}125 \ \ +5{,}125\,][D\,] + [+3{,}5 \ \ -3{,}5\,][F\,]$$

$$+ \begin{bmatrix} +6{,}125 & -6{,}125 \\ -6{,}125 & +6{,}125 \end{bmatrix} AD + \begin{bmatrix} -5{,}125 & +5{,}125 \\ +5{,}125 & -5{,}125 \end{bmatrix} DF$$

3-4 : Modélisation en représentation qualitative

3-4-2 : Application : recherche des modalités conduisant à une réponse optimale

Exemple : exercice n°5 , 7ème question b)-

$E_A = 10{,}25$	$E_D = 5{,}125$	$E_F = -3{,}5$
$E_{AD} = 6{,}125$	$E_{DF} = -5{,}125$	$\bar{y}_c = 116$

La réponse maximale sera obtenue avec les modalités suivantes :

- A et D : niveau haut (E_A et E_D positifs)

- F : niveau bas (E_F négatif)

Réponse prédite :

$$\hat{y} = 116 + 10{,}25 + 5{,}125 + 3{,}5 + 6{,}125 + 5{,}125 = \mathbf{146{,}125}$$

- On retrouve la réponse prédite avec le modèle quantitatif linéaire
- On pouvait « deviner » ces modalités optimales en repérant les réponses mesurées les plus élevées : essais n°10, 12, 14 et 16

Conclusion sur les plans 2^n

Ces plans sont efficaces, simples à analyser et à interpréter. Ils ont également des vertus pédagogiques en amenant l'expérimentateur à sortir de la stratégie : « un seul facteur variable étudié à la fois »

Cependant :

❶ Quand le nombre de facteurs augmente : expérimentations coûteuses en temps et en argent -→ plans fractionnaires

❷ Certaines combinaisons de facteurs peuvent être irréalisables : par exemple un composé peut précipiter lorsqu'on utilise un pH et une température incompatibles.

❸ Deux niveaux par facteur peuvent être insuffisants : par exemple il faut plus de 2 concentrations pour faire une courbe d'étalonnage…-→ plans avec plus de 2 niveaux