

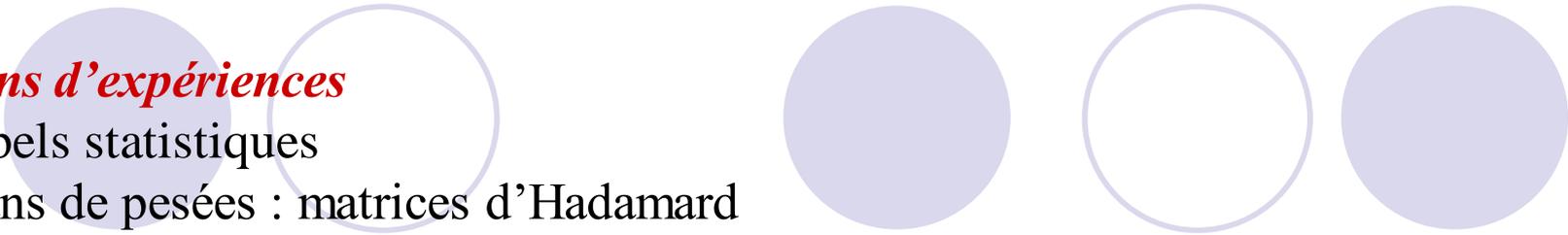
*UNIVERSITE MOHAMMED V-AGDAL
FACULTE DES SCIENCES
DEPARTEMENT DE CHIMIE
RABAT*

Cours de Master Chimie

Chimiométrie

A 3D plot with three axes (x, y, z) extending from a central origin. The axes are colored in a gradient from purple to blue. The plot is set on a grid of orange and yellow squares. A green, curved surface is plotted on the grid, resembling a paraboloid or a similar mathematical function. The word 'Chimiométrie' is written in white, italicized font across the center of the plot.

Pr R. BCHITOU



A- Plans d'expériences

I- Rappels statistiques

II - Plans de pesées : matrices d'Hadamard

III - Modélisation des essais : Régression multiple

IV - Matrice d'Hadamard

Le criblage des facteurs expérimentaux, : classement hiérarchisé des facteurs

V - Plans factoriels complets FFD

VI - Surface de réponse Plans composites CCD

B- Analyse de données

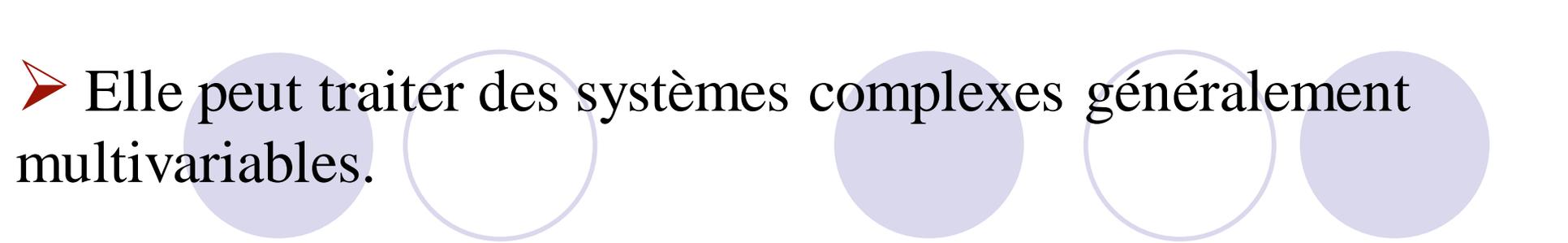
- Analyse en Composantes Principales

- Analyse Factorielle des Correspondances

Définition de la Chimométrie

La Chimométrie est un outil utilisé afin d'extraire de l'information pertinente et utile à partir des données physicochimiques mesurées ou connues brutes. Aussi, on l'appelle souvent en chimie « Analyse multivariable ».

Le terme chimométrie vient de l'anglais «chemometrics », jeune discipline associant initialement l'analyse des données et la chimie analytique.

- 
- Elle peut traiter des systèmes complexes généralement multivariables.
 - Elle recouvre l'ensemble des applications de la chimie, de la physique, des sciences de la vie ...
 - Elle est basée sur des règles mathématiques strictes et des démarches rigoureuses de la part de l'expérimentateur.
 - La méthodologie repose sur la modélisation et l'optimisation:
 - *l'exploitation d'un modèle* de comportement à l'aide des outils statistiques.
 - *la construction.*

But de la méthodologie

✓ Organiser mathématiquement les conditions expérimentales pour choisir les plus informatives permettant de minimiser le nombre d'expériences tout en maximisant l'information obtenue

L'applications de la **Chimométrie** dans le domaine de la chimie analytique consiste à :

✓ **modéliser** les variations d'un certain nombre de variables Y, que nous appellerons réponses (l'analyse chimique par exemple) en fonction d'autres variables X appelées variables mesurables (mesure de paramètres physico-chimiques par exemple);

✓ **optimisation** des procédés chimiques.

On distingue 2 opérations :

- L'étalonnage en laboratoire où toutes les mesures de variables doivent être réalisées et où le modèle est calculé.

- La prédiction : utilisation courante où seules les variables X sont mesurées, les autres Y, étant calculées à l'aide du modèle.

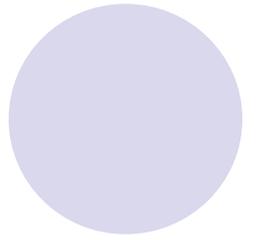
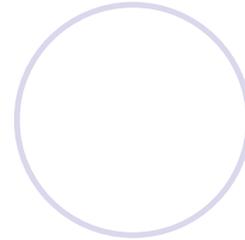
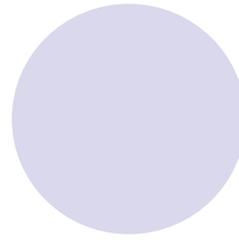
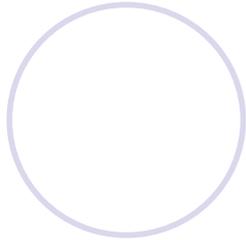
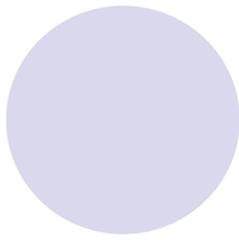
Intérêt de la méthodologie

- Définir la région du domaine expérimental où la réponse satisfait une contrainte.
- Optimiser l'organisation des essais en déterminant les facteurs influents à partir d'un modèle.
- Obtenir la meilleure précision possible sur la modélisation des résultats.

Avantages des plans d'expériences

Les plans d'expériences permettent une diminution considérable du nombre d'essais et une interprétation rapide et sans équivoque. Ils fournissent des résultats faciles à présenter.

- possibilité d'étudier un très grand nombre de facteurs
- détection des interactions éventuelles
- modélisation aisée des résultats
- détermination des résultats avec une bonne précision.



Introduction
à la pratique des
Plans d'expériences

Plans d 'expériences : la Méthode

La Méthode

" Toute la méthode consiste dans l'ordre et la disposition des choses dans lesquelles il faut tourner le regard de l'esprit, pour découvrir quelque vérité "



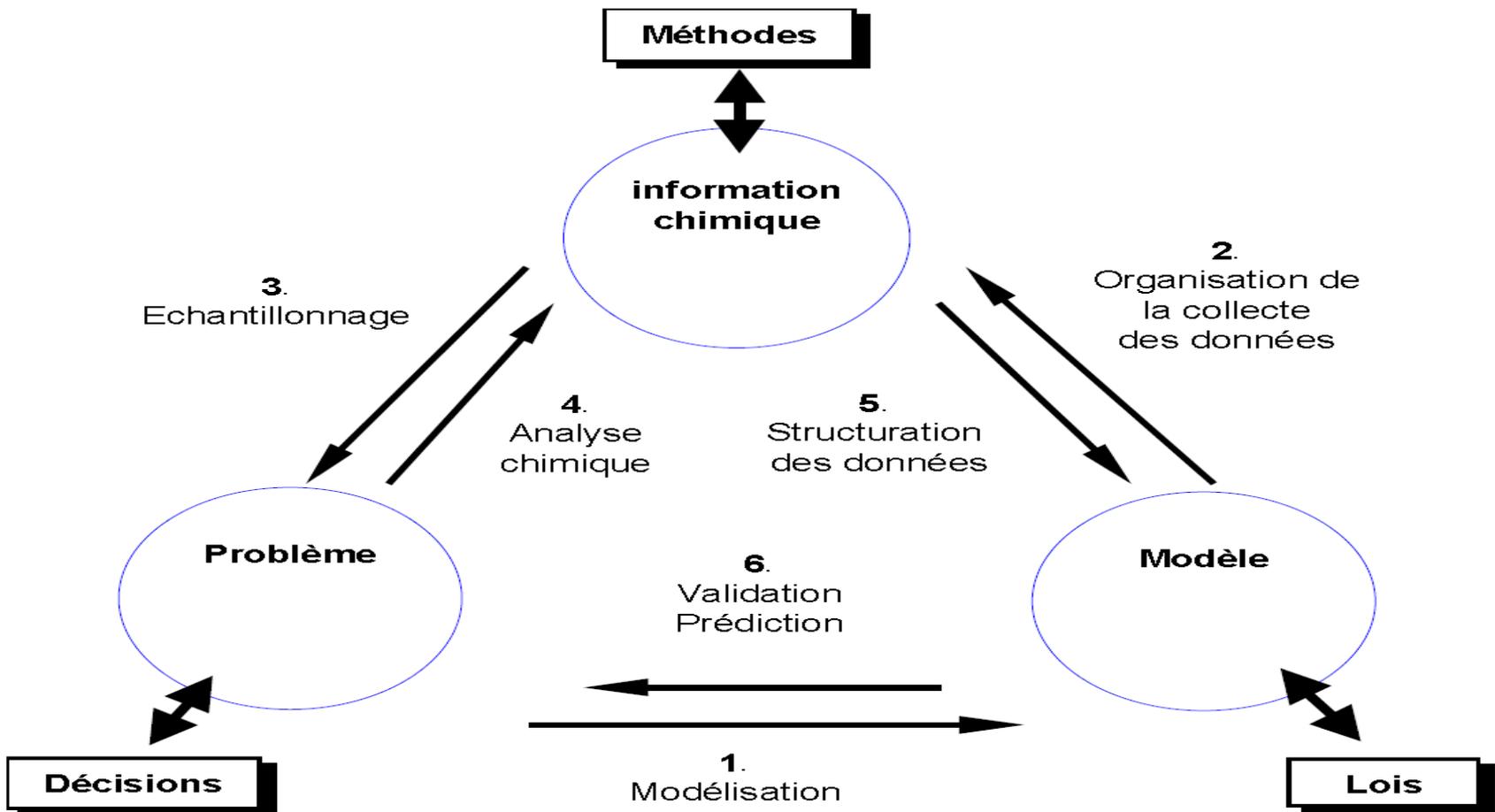


Figure 1 : Démarche de l'Analyticien : place de l'organisation de la collecte des données

Chimiométrie

Données Issues
d'essais structurés ?

OUI

Plans d'Expériences

NON

Analyse de Données

Expériences

Entrées

Sorties

Causes

Effets

Facteurs

Réponses



X_1

X_2

X_3



Y_1

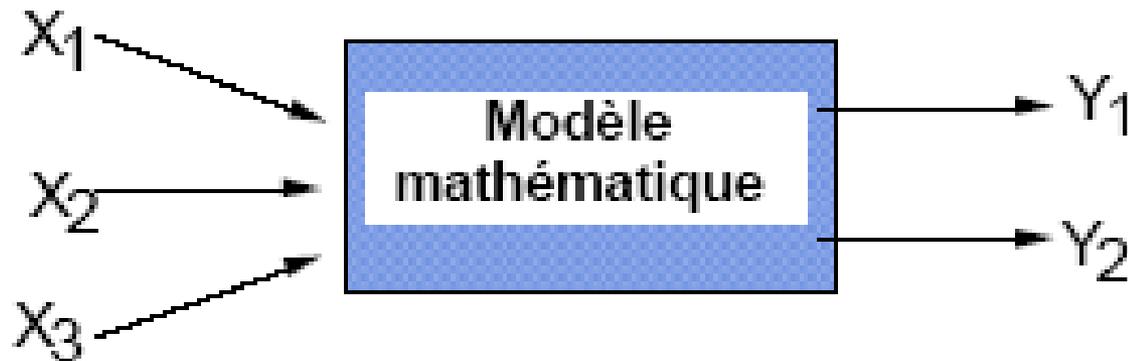
Y_2



Plans d 'expérience

Facteurs

Réponses



Variation
Raisonnée
des Facteurs

Calcul des Effets
sur les réponses

MATRICES D'EXPERIENCES et METHODOLOGIE

Les objectifs :

- **Le criblage des facteurs** : classement hiérarchisé des facteurs;
- **Les études quantitatives des facteurs** : quantification des influences principales et des synergies éventuelles;
- **Les études quantitatives des réponses** : modélisation prévisionnelle du phénomène étudié;
- **L'optimisation** : déterminer un ou plusieurs points de fonctionnement optimaux .

Les Modèles : polynômes

Soit X un facteur quantitatif. Il peut être représenté par un polynôme dans les modèles :

Premier degré \Rightarrow Y (réponse) = $b_0 + b_1 X_1 + b_2 X_2$

Effet principal

Synergique \Rightarrow Y (réponse) = $b_0 + b_1 X_1 + b_2 X_2 + b_{12} X_1 X_2$

Effet d'interaction

Quadratique \Rightarrow Y (réponse) = $b_0 + b_1 X_1 + b_2 X_2 + b_{12} X_1 X_2 + b_{11} X_1^2 + b_{22} X_2^2$

Surface de réponse

« Courbure »

Les MATRICES

premier degré

$$Y (\text{réponse}) = b_0 + b_1 X_1 + b_2 X_2$$

MATRICES

Hadamard
(Plackett-Burman)

synergique

$$Y (\text{réponse}) = b_0 + b_1 X_1 + b_2 X_2 + b_{12} X_1 X_2$$

Factorielles

quadratique

$$Y (\text{réponse}) = b_0 + b_1 X_1 + b_2 X_2 + b_{12} X_1 X_2 + b_{11} X_1^2 + b_{22} X_2^2$$

**Composite
(Doehlert)**

Représentation matricielle

$$\begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ 1 & x_3 \\ 1 & x_4 \\ \dots & \dots \\ 1 & x_n \end{pmatrix} * \begin{pmatrix} b_0 \\ b_1 \end{pmatrix} = \begin{pmatrix} \hat{y}_1 \\ \hat{y}_2 \\ \hat{y}_3 \\ \hat{y}_4 \\ \dots \\ \hat{y}_n \end{pmatrix}$$

Vecteur des coefficients : **B**

Vecteur de la réponse prédite par le modèle :

\hat{Y}

Matrice du modèle
X

$$\mathbf{X} \mathbf{B} = \hat{\mathbf{Y}}$$

Rappel de quelques définitions concernant le plan d'expériences

- **Réponse** : Le résultat mesuré d'une étude.

A chaque point du domaine d'étude correspond une réponse. L'ensemble des réponses forme la surface de réponse.

- **Variables explicatives et notion d'interaction** : Les variables explicatives d'une étude sont les paramètres susceptibles de modifier les réponses de cette étude.

Si l'effet d'une variable explicative dépend du niveau d'une autre variable explicative, on dit qu'il y a interaction entre ces deux variables explicatives.

Niveaux d'une variable explicative :

Les niveaux d'une variable explicative sont les différents états que peut prendre cette variable explicative

Notion d'effet significatif :

L'effet d'une variable explicative sur la réponse y s'obtient en comparant les deux résultats de mesure y_1 et y_2 de réponse, mesurée lorsque la variable explicative passe d'un niveau (0) à un niveau (+). Si l'écart entre y_1 et y_2 est important on dit que le facteur est influent ou significatif.

Variables codées et variables naturelles :

Les variables naturelles x_i sont les valeurs qui correspondent à chaque niveau d'une variable explicative. Pour comparer les effets des variables naturelles sur la réponse, il est nécessaire de les remplacer par les variables codées X_i qui sont sans unité.

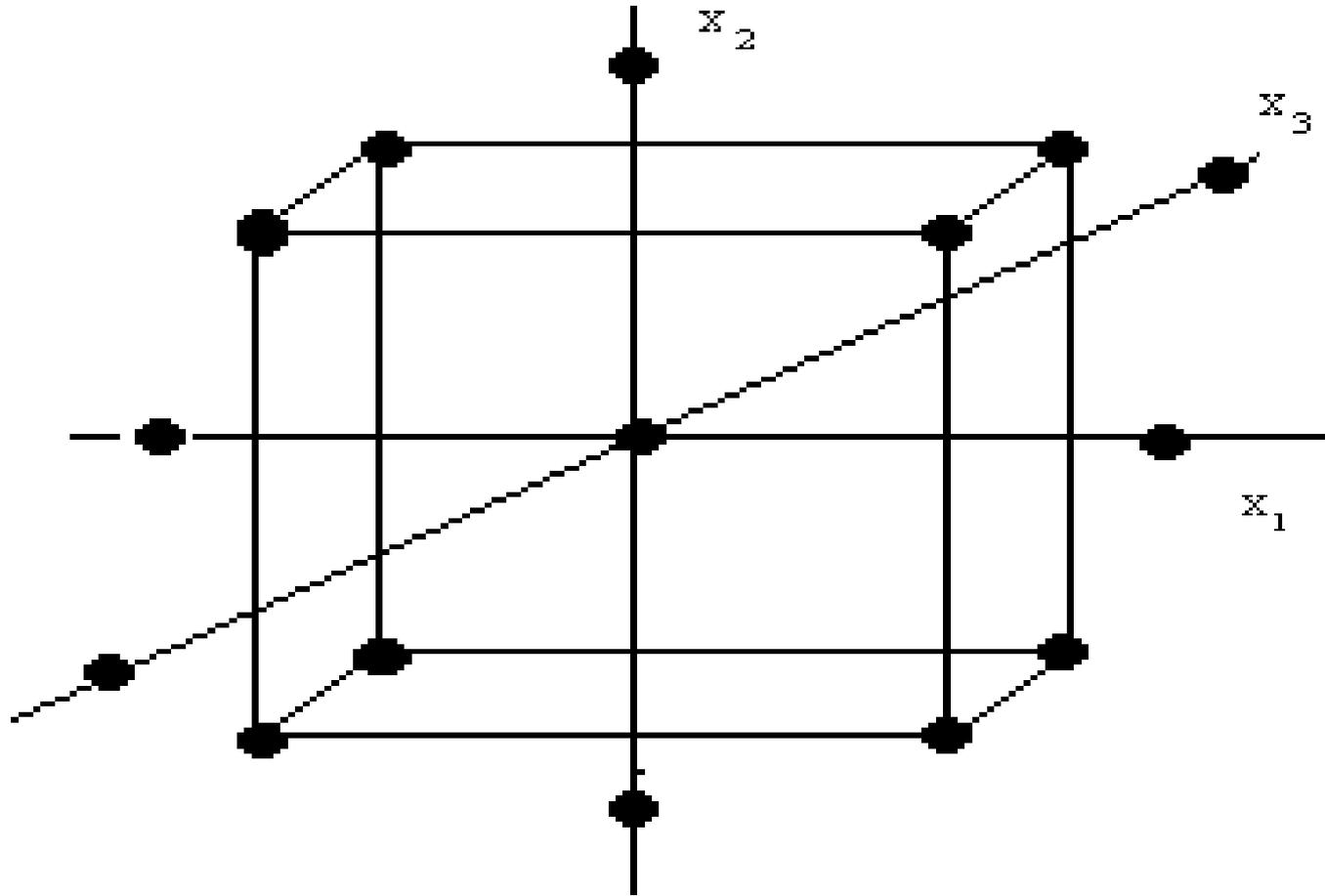
● **Matrice d'expérience :**

Matrice d'expérience est un tableau de n lignes et k colonnes, regroupant les conditions expérimentales d'un plan d'expériences. n et k correspondent respectivement au nombre d'expériences et au nombre des variables codées.

● **Courbes d'isoréponses :**

Après la détermination du modèle et la vérification de sa validité, les courbes d'isoréponses peuvent être tracé à l'intérieur du domaine expérimental. Ces courbes représentent des plans pour surfaces de réponse c'est à dire la représentation graphique des résultats (modèle estimé) pour pouvoir en tirer des optimums.

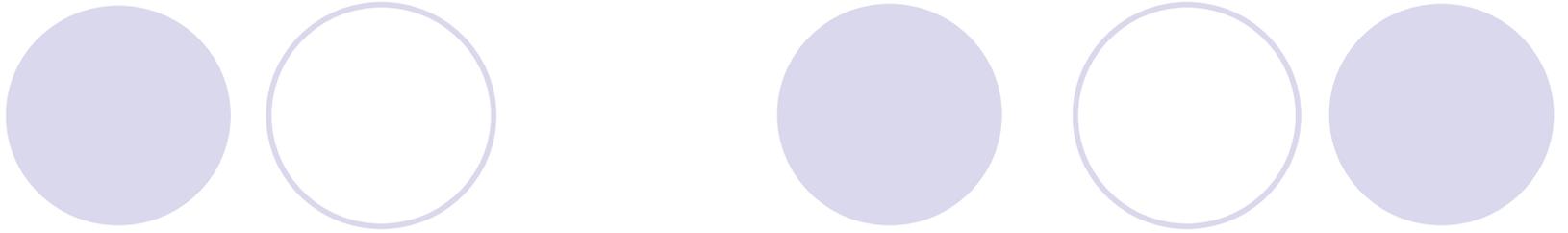
● La représentation dans 3^3 de ce plan est donnée dans la figure suivante :



Représentation d'un plan composite centré en 3 dimensions

CHAPITRE I

Rappels statistiques



I/ Régression

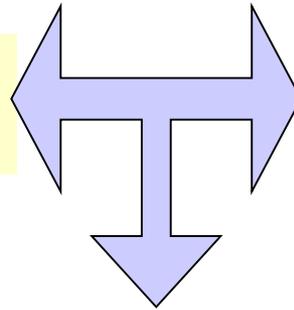
I-1/ Définition de la régression

La régression est une méthode très utilisée en estimation.

A partir d'un ensemble de valeurs expérimentales représentées par des points sur un graphique , on cherche à calculer la courbe qui reproduit le mieux les variations de la grandeur à étudier, c'est-à-dire celle qui passe par tous les points ou le plus proche possible.

I-2/ Régression linéaire

Si X représente une **teneur connue** en analyte



Y représente le **résultat observé**,

On peut disposer de n couples $[X_i, y_i]$ pour deux variables X et Y que l'on suppose liées : à chaque valeur de X est associée une valeur de Y avec la relation :

$$(Y = \beta_1 X + \beta_0)$$

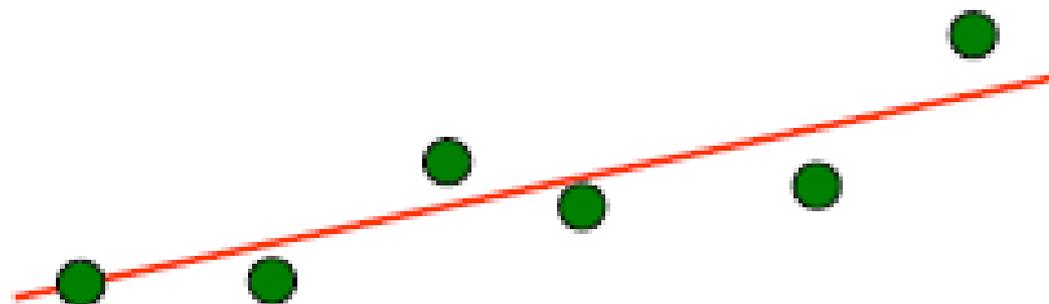
Mais, **expérimentalement**, à chaque valeur x_i de X , on obtient une valeur y_i entachée de l'erreur expérimentale ϵ_i . On a en réalité :

$$y_i = \beta_1 x_i + \beta_0 + \epsilon_i$$

A cause de cette erreur ε_i associée à chaque couple $[x_i, y_i]$, si on représente graphiquement y_i en fonction de x_i ,

➡ on ne va pas obtenir des points "idéalement alignés",

➡ mais un «nuage» de points plus ou moins écartés de cette droite idéale.



I- 3/ Analyse quantitative et Étalonnage

Les données sont toujours en nombre limité

elles ne représentent donc qu'un échantillon de la population de toutes les mesures de la teneur en analyse de l'étalon que l'on pourrait effectuer.

X représente une teneur connue en analyse

Y représente le résultat observé,

la relation linéaire postulée devient : $Y = b_1 X + b_0$

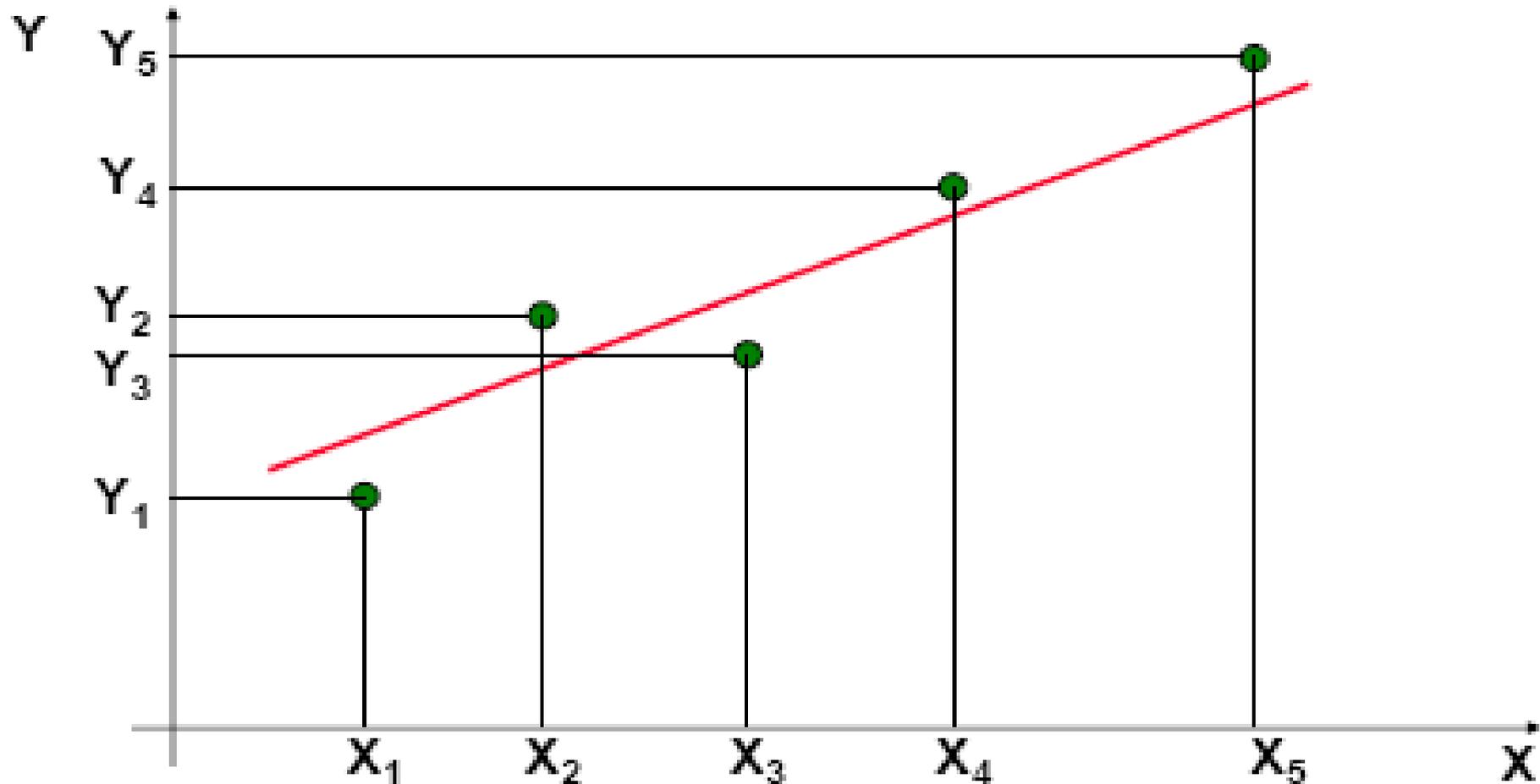
Avec uniquement une « estimation » des coefficients b_0 et b_1 du modèle postulé.

$$Y = \beta_1 X + \beta_0$$

Modèle linéaire

Avec **une seule variable** X le modèle s'écrit :

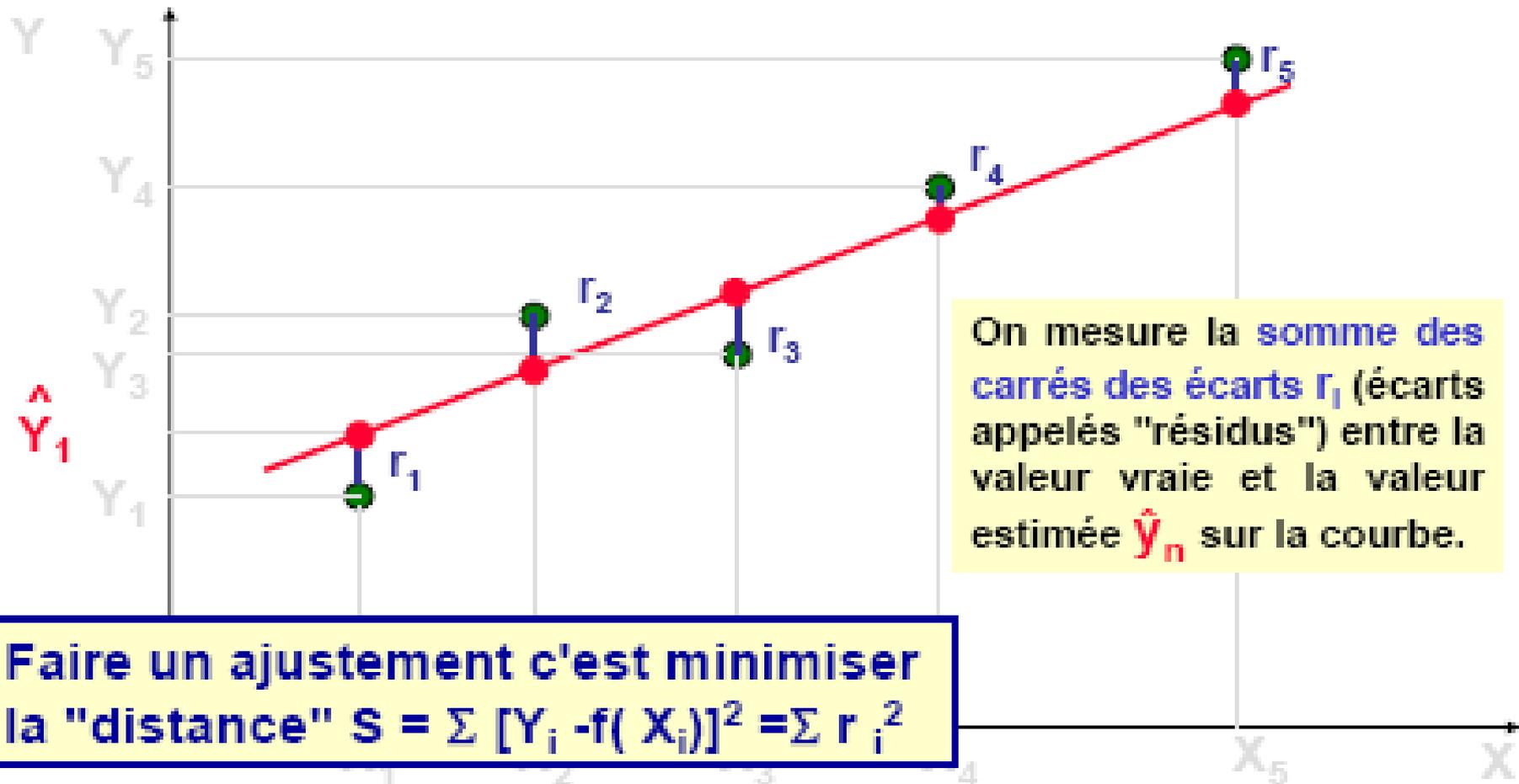
$$Y = \beta_0 + \beta_1 X + r$$



Ajustement linéaire

Avec **une seule variable** X le modèle s'écrit :

$$Y = \beta_0 + \beta_1 X + r$$



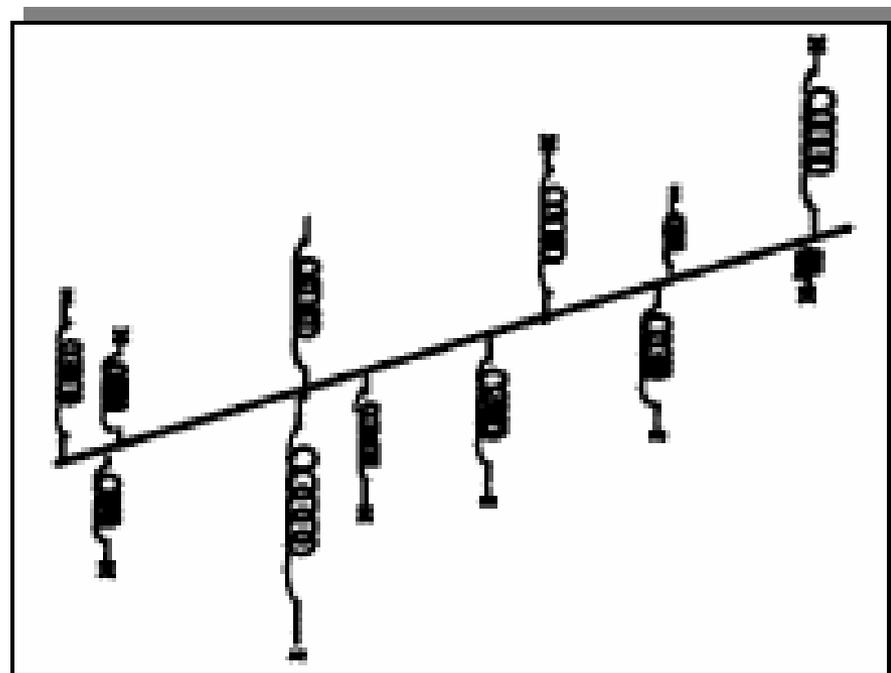
Droite des moindres carrés et efficacité d'un ajustement

La somme S des carrés des écarts entre les valeurs expérimentales et les valeurs calculées par le modèle s'écrit :

$S = \sum [y_i - (\beta_0 + \beta_1 x_i)]^2$ est une fonction de β_0 et β_1 .

Pour minimiser S ,
il suffit d'annuler les
dérivées partielles
de S par rapport à
 β_0 et à β_1 :

$$\frac{\partial S}{\partial \beta_0} = \frac{\partial S}{\partial \beta_1} = 0.$$



Estimation des coefficients

Dans ce système les β_i sont les inconnues que nous devons estimer :
(b_i est l'estimation calculée de β_i).

1. Au sens des moindres carrés (résolution algébrique) :

$$\hat{\beta}_1 = b_1 = \frac{\sum (y_i - \bar{y})(x_i - \bar{x})}{\sum (x_i - \bar{x})^2} \quad \text{et} \quad \hat{\beta}_0 = b_0 = \bar{y} - b_1 \bar{x}$$

2. Au sens des moindres carrés (résolution matricielle) :

Expression matricielle de la régression

$$\begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ 1 & x_3 \\ 1 & x_4 \\ \dots & \dots \\ 1 & x_n \end{pmatrix} * \begin{pmatrix} b_0 \\ b_1 \end{pmatrix} + \begin{pmatrix} r_1 \\ r_2 \\ r_3 \\ r_4 \\ \dots \\ r_n \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ \dots \\ y_n \end{pmatrix}$$

\hat{y}_1
 \hat{y}_2
 \hat{y}_3
 \hat{y}_4
 \dots
 \hat{y}_n

Vecteur des résidus : r

Vecteur de la réponse expérimentale

$XB + r = Y$

Analyse de la régression linéaire

Les variations observées pour Y sont-elle dues globalement aux variations de X ?

Quelle confiance peut en avoir :

$$\hat{Y}_i = b_1 X_i + b_0$$

➔ D'une part globalement pour la régression :

- Analyse de variance/ coefficients
- Examen des résidus
- Manque d'ajustement (Lack of fit)

➔ D'autre part individuellement pour les estimateurs :

- Simplification du modèle
- Pertinence quadratique global

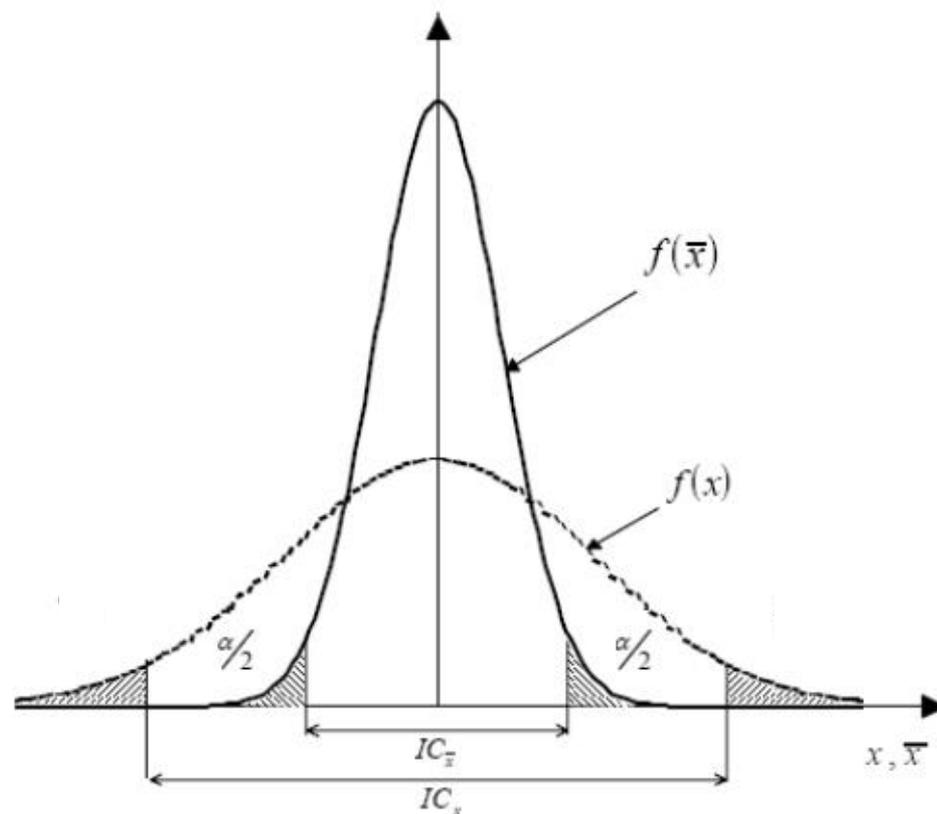
Propriétés de la loi normale

• Moyenne :

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$$

• Variance :

$$\hat{\sigma}^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$$



Le graphe de la Loi Normale est caractérisé par :



Une courbe en cloche asymptotique à l'axe des x , dont le maximum est pour $x = \bar{x}$,

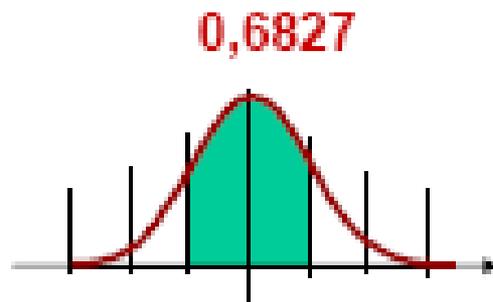


Une symétrie par rapport à l'axe $x = \bar{x}$,

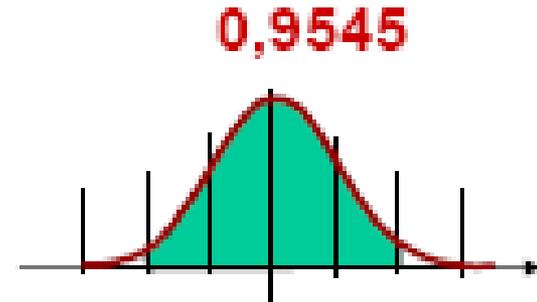


Deux points d'inflexion à une distance de \bar{x} égale à σ .

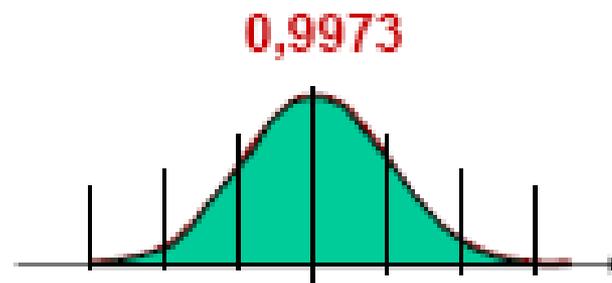
Propriétés de la loi Normale



Probabilité = 68,27% pour que x soit compris dans l'intervalle $\bar{x} \pm 1 \sigma$



Probabilité = 95,45% pour que x soit compris dans l'intervalle $\bar{x} \pm 2 \sigma$



Probabilité = 99,73% pour que x soit compris dans l'intervalle $\bar{x} \pm 3 \sigma$

Caractéristiques de l'erreur expérimentale r_i

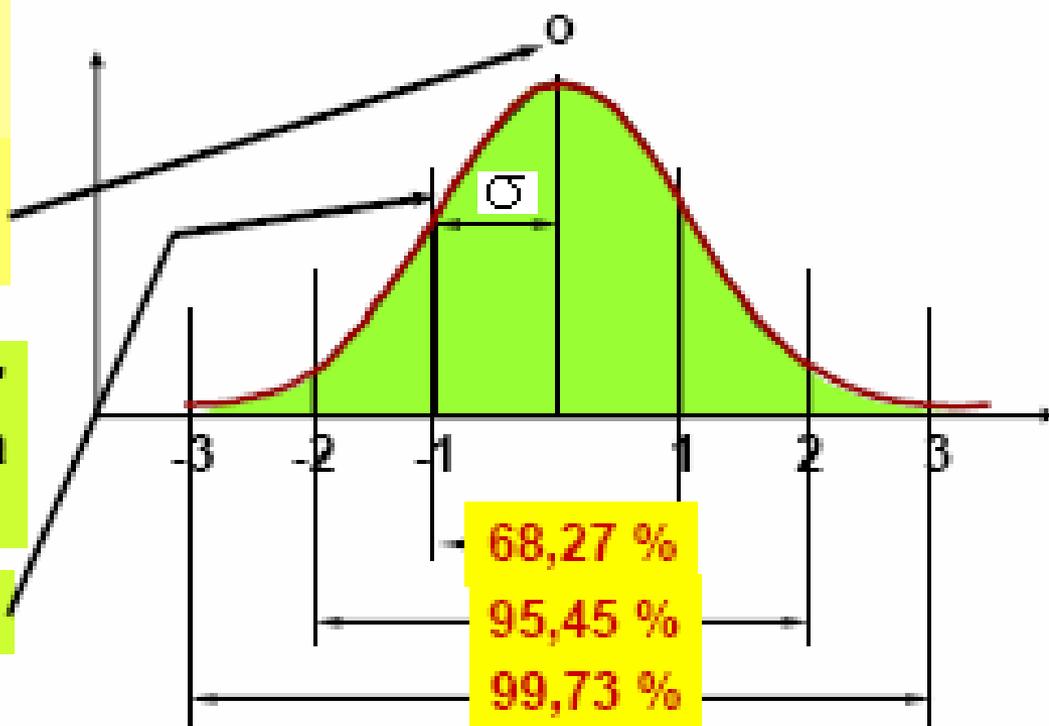
Distribution de Gauss centrée sur zéro
(échelle des abscisses en unités d'écart-type)

En **moyenne**,
l'erreur **est nulle** :

l'espérance
mathématique $E(r_i) = 0$.

La **dispersion** de « r_i »
est mesurée par sa
variance : $\text{var}(r_i) = \sigma^2$

ou par l'écart-type σ .



Significativité des coefficients

b_1 estimation de β_1 de
moyenne β_1 et de
variance $\text{var}(b_1)$

b_0 estimation de β_0 de
moyenne β_0 et de
variance $\text{var}(b_0)$

Comme la variable Y qui intervient dans ces calculs est une **variable aléatoire** de variance $\sigma^2_{\text{exp.}}$

cette dispersion va se répercuter sur
les variances de b_0 et b_1 .

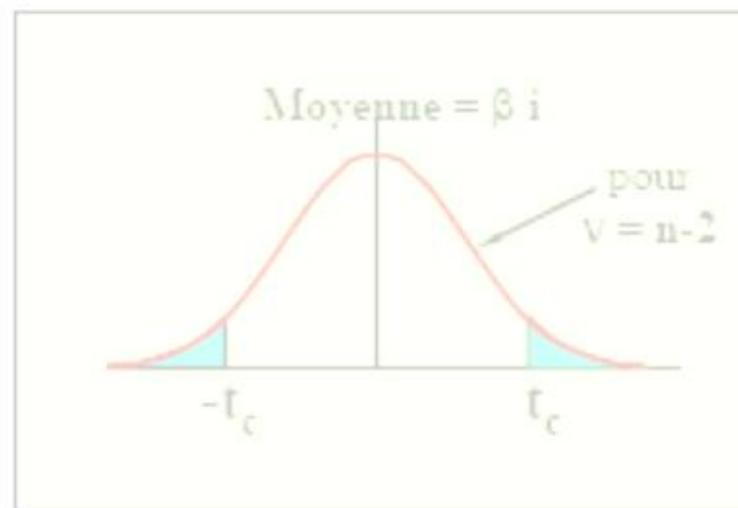


$$\text{Var}(b_1) = \frac{\sigma^2_{\text{exp.}}}{\sum(x_i - \bar{X})^2}$$

$$\text{Var}(b_0) = \sigma^2 \left[\frac{1}{n} + \frac{\bar{X}^2}{\sum(x_i - \bar{X})^2} \right]$$

Significativité des coefficients

Le coefficient b_i est distribué selon une distribution de Student de moyenne β_i , d'écart-type $e.t.(b_i)$ et $(n-2)$ degrés de liberté.



Intervalle de confiance pour b_i :

$$b_i \pm t_c e.t.(b_i)$$

Significativité des coefficients

la différence $b_i - \beta_i^0$ suit une statistique de Student à $v = (n-2)$ degrés de liberté avec :

$$t = \frac{b_i - \beta_i^0}{\text{é.type } (b_i)}$$

La **significativité** va être **déterminée en prenant** $\beta_i^0 = 0$ d'où :

$$t = \frac{b_i}{\text{é.type } (b_i)}$$

Ceci va être vérifié par la Loi de student (Voir Tableau)

Résolution matricielle



Matrice de variance-covariance des coefficients

C'est une matrice où **les variances** sont disposées **sur la diagonale** et **les covariances de part et d'autre** de cette diagonale (matrice carrée symétrique) :

Variance-covariance

$$\hat{\sigma}_x^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$$

Variance de x

$$\hat{\sigma}_y^2 = \frac{1}{n-1} \sum_{i=1}^n (y_i - \bar{y})^2$$

Variance de y

$$\hat{\sigma}_{x,y} = \frac{1}{n-1} \sum_{i=1}^n (y_i - \bar{y})(x_i - \bar{x})$$

Covariance xy

Matrice de variance-covariance des coefficients

$$\text{var}(\mathbf{B}) = \begin{bmatrix} \text{var}(b_0) & \\ \text{cov}(b_1, b_0) & \text{var}(b_1) \end{bmatrix}$$

$$\text{Var}(\mathbf{B}) = \sigma^2 (\mathbf{X}'\mathbf{X})^{-1}$$

Variance expérimentale

Conséquence:

le choix des points expérimentaux conditionne la qualité de l'estimation

→ la meilleure estimation consiste à annuler la covariance et minimiser les variances sur les coefficients

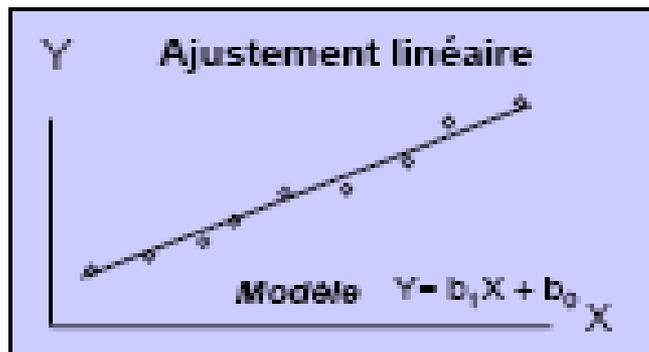
→ **Plan d'expériences**

II/ Régression multiple

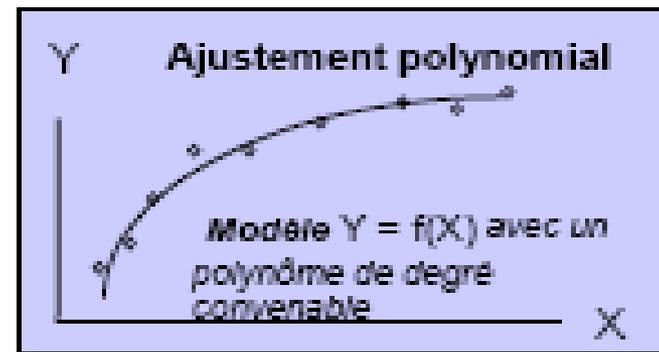
Modélisation

Modéliser : utiliser des données expérimentales pour **prévoir une information quantitative** inconnue Y à partir de mesures de X via une certaine « **fonction mathématique** » :

Le modèle mathématique postulé **peut être** :



Une droite si Y varie linéairement avec X .



Sinon un polynôme de degré convenable.

Régression multiple

Exemple : Etude de la stabilité d'une suspension

| Facteurs expérimentaux | Variables | Domaine Expérimental | |
|------------------------------|-----------|----------------------|---------|
| | | Minimum | Maximum |
| Tensio-actif : Mouillant 1 | M_1 | 20 g/l | 40 g/l |
| Tensio-actif : Mouillant 2 | M_2 | 5 " | 15 " |
| Epaississant : Structurant 1 | S_1 | 5 " | 20 " |
| Epaississant : Structurant 2 | S_2 | 0 " | 10 " |

Réponse étudiée :
Y = % de séparation de la suspension en deux phases

Tableau des résultats

Plan d'expérimentation

Réponse

| Essais | M ₁ | M ₂ | S ₁ | S ₂ | Y |
|--------|----------------|----------------|----------------|----------------|------|
| 1 | 40 | 15 | 20 | 0 | 10 |
| 2 | 40 | 15 | 5 | 10 | 16 |
| 3 | 40 | 5 | 20 | 0 | 15 |
| 4 | 20 | 15 | 5 | 0 | 38,7 |
| 5 | 40 | 5 | 5 | 10 | 30,5 |
| 6 | 20 | 5 | 20 | 10 | 18 |
| 7 | 20 | 15 | 20 | 10 | 13 |
| 8 | 20 | 5 | 5 | 0 | 32 |

Modèle postulé

$$Y = a_0 + a_1 M_1 + a_2 M_2 + a_3 S_1 + a_4 S_2$$

Codage des Variables

Variables naturelles U et
Variables codées X

U_i

Toutes les variables naturelles ont leurs propres plage de variation qui dépendent des unités

Les coefficients du modèle **ne peuvent pas** se comparer directement

Interprétation difficile

X_i

Toutes les variables codées ont la même plage de variation entre -1 et +1 indépendante des unités

Les coefficients du modèle **peuvent** se comparer directement

Interprétation facile

Codage des Variables

Variables naturelles U et
Variables codées X

= **moyenne** des valeurs maximum et minimum que peut prendre la variable U_j .

U_{ij} = valeur de la variable naturelle j à l'expérience i.

$$X_{ij} = \frac{U_{ij} - U_j^0}{\Delta U_j}$$

U_j^0 = valeur de la variable naturelle j au **centre du domaine**.

ΔU_j = **pas de variation** de la variable naturelle j.

X_{ij} = valeur de la variable codée j pour l'expérience i.

Réciproquement :

$$U_{ij} = U_j^0 + X_{ij} \cdot \Delta U_j$$

= **demi étendue** : moitié de l'écart entre la valeur maximum et la valeur minimum que peut prendre la variable U_j .

Valeurs codées

U_1 est le **mouillant 1** qui varie entre
20 et 40 g/l

centre du domaine U_1
 $= (20+40)/2 = 30 \text{ g/l}$

Pas de variation ΔU_1
 $= (40-20)/2 = 10 \text{ g/l}$

à 20 g/l: $X_{i1} = (20 - 30)/10 = -1$
à 40 g/l: $X_{i1} = (40 - 30)/10 = 1$

U_4 est le **structurant 2** qui varie entre
0 et 10 g/l

centre du domaine U_4^0
 $= (0+10)/2 = 5 \text{ g/l}$

Pas de variation ΔU_4
 $= (10-0)/2 = 5 \text{ g/l}$

à 0 g/l : $X_{i4} = (0-5)/5 = -1$
à 10 g/l : $X_{i4} = (10-5)/5 = 1$

Régression avec des variables codées

Matrice d'expériences

| Essais | X ₁ | X ₂ | X ₃ | X ₄ |
|--------|----------------|----------------|----------------|----------------|
| 1 | +1 | +1 | +1 | -1 |
| 2 | +1 | +1 | -1 | +1 |
| 3 | +1 | -1 | +1 | -1 |
| 4 | -1 | +1 | -1 | -1 |
| 5 | +1 | -1 | -1 | +1 |
| 6 | -1 | -1 | +1 | +1 |
| 7 | -1 | +1 | +1 | +1 |
| 8 | -1 | -1 | -1 | -1 |

Modèle postulé

$$Y = b_0 + b_1 X_1 + b_2 X_2 + b_3 X_3 + b_4 X_4$$

Statistiques des coefficients

| | Coefficients b_1 | Erreur | Statist. t | Probab. (signif.) | Seuil de conf. 95% | |
|-----------|-----------------------|--------|------------|----------------------|--------------------|----------|
| | | | | | Lim.inf. | Lim.sup. |
| Constante | 21,65 | 2,17 | 9,98 | 0,002 | 14,75 | 28,55 |
| X_1 | -3,775 | 2,17 | -1,74 | 0,180 | -10,68 | 3,13 |
| X_2 | -2,225 | 2,17 | -1,03 | 0,380 | -9,13 | 4,68 |
| X_3 | -7,65 | 2,17 | -3,53 | 0,040 | -14,55 | -0,75 |
| X_4 | -2,275 | 2,17 | -1,05 | 0,370 | -9,18 | 4,63 |

$$t_{(\nu=3, \alpha=0.05)} = 3.18$$

$$\text{Intervalle de confiance} \pm 3.18 \cdot 2.17 = \pm 6.9$$

Les Outils pour l'Analyse des Résultats

Etude des Résidus

Reprenons l'étude de la stabilité d'une suspension avec
 $Y = \%$ de séparation de la suspension en deux phases

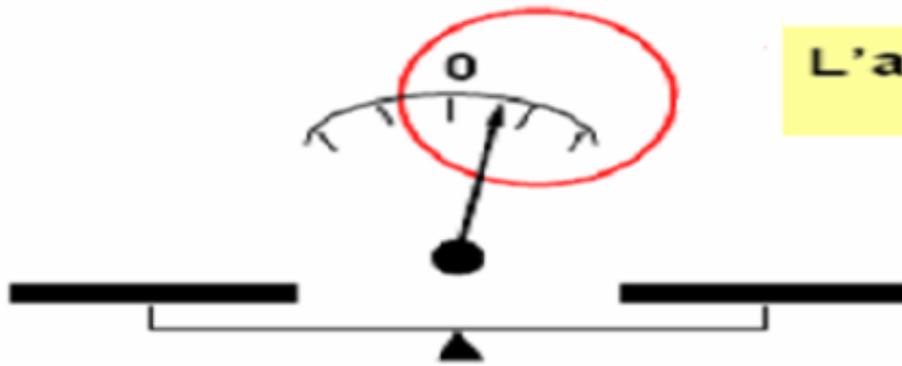
| Essais | Y_{exp} $y(i)$ | $Y_{\text{calc.}}$ $y_{\text{pred}}(i)$ | $Y_{\text{exp.}} - Y_{\text{calc.}}$ = résidus | Résidus normés | Différence Normée = différence/écart-type |
|--------|----------------------------|--|---|-------------------|---|
| 1 | 10 | 10.275 | -0.275 | -0.045 | Écart-type résiduel = 6,136 $r_1 = -0,275$ $r_1 \text{ normé} =$ $-0,275/6,136 = -0,0448$ |
| 2 | 16 | 21.025 | -5.025 | -0.819 | |
| 3 | 15 | 14.725 | 0.275 | 0.045 | |
| 4 | 38.7 | 33.125 | 5.575 | 0.909 | |
| 5 | 30.5 | 25.475 | 5.025 | 0.819 | |
| 6 | 18 | 17.725 | 0.275 | 0.045 | |
| 7 | 13 | 13.275 | -0.275 | -0.045 | |
| 8 | 32 | 37.575 | -5.575 | -0.909 | |

Chapitre II

Plans de pesée

Expérience

Balance



L'aiguille **n'est pas** sur le zéro

Nous convenons d'appeler "expérience" **le fait de lire la position de l'aiguille** sur le cadran.

Comme pour toute expérience, **il existe une erreur expérimentale** :



erreur de parallaxe, épaisseurs relatives de la pointe de l'aiguille et des traits de graduation etc.

Efficacité

Etre efficace c'est rendre minimum :

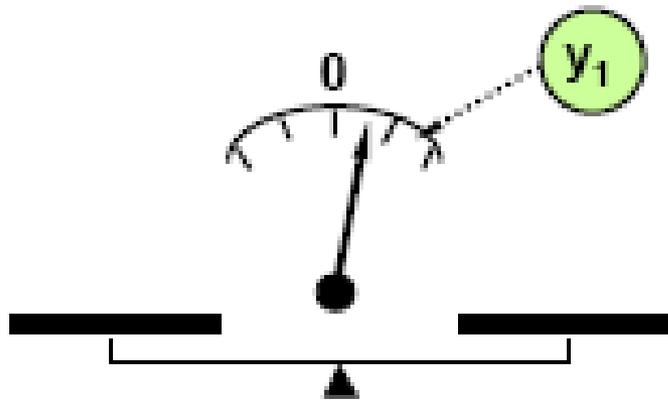
non pas l'erreur de lecture

MAIS

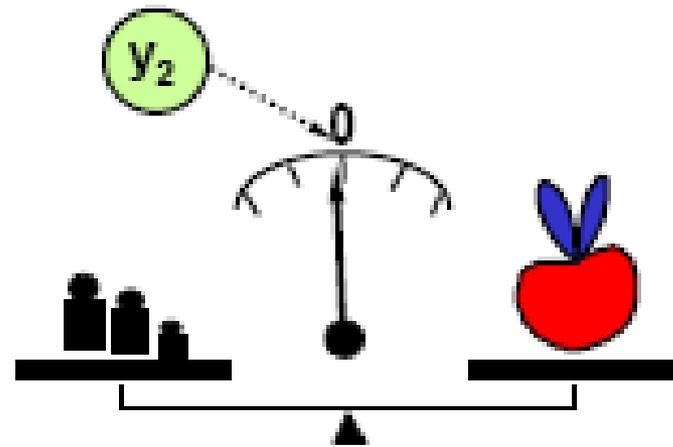
sa répercussion sur l'estimation des masses,

nous admettrons que l'expérimentateur prend toutes les précautions utiles pour la rendre minimale

(sur le calcul). En d'autres termes c'est cerner au plus près la masse vraie des objets à peser.



Lecture à vide



Lecture avec la pomme

$$\text{Masse de la pomme} = y_2 - y_1$$

Chaque pesée, matérialisée par la lecture Y_i (réponse), est :

indépendante des autres lectures

une **variable aléatoire** car Y_i est une **mesure expérimentale**.

Y_i est en réalité la **somme de deux valeurs** :

$$Y_i = \eta_i + e_i$$

η_i = valeur **vraie**

e_i = **erreur** de mesure aléatoire

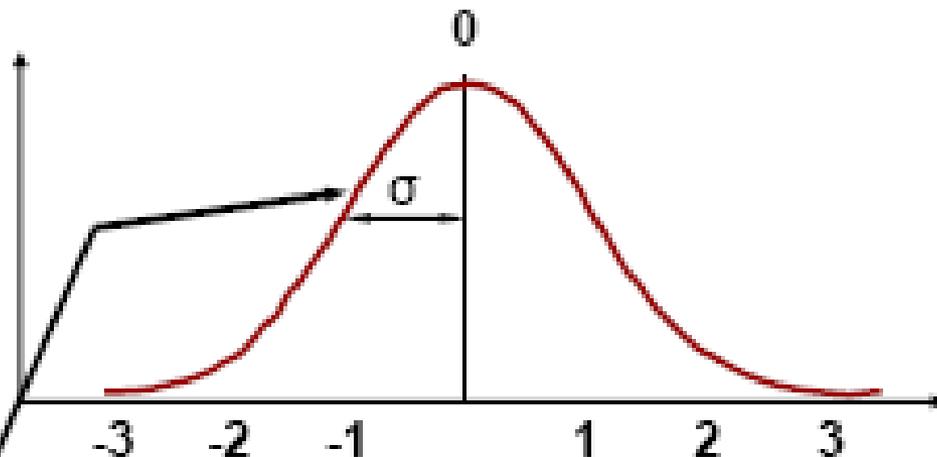
Caractéristiques de l'erreur expérimentale r_i

Distribution de Gauss centrée sur zéro

En **moyenne**,
l'erreur **est nulle**

Sa **dispersion** est
mesurée par sa
variance : **$\text{var}(r_i) = \sigma^2$**

ou par l'écart-type σ .



Qualité de l'estimation de la masse de la pomme

Revenons à l'estimation de la Masse de la pomme $m = y_2 - y_1$

$$\begin{aligned}\text{var}(m) &= \text{var}(Y_2 - Y_1) = \text{var}(Y_2) + \text{var}(Y_1) \\ &= \text{var}(\eta_2 + e_2) + \text{var}(\eta_1 + e_1) \\ &= \text{var}(\eta_2) + \text{var}(e_2) + \text{var}(\eta_1) + \text{var}(e_1)\end{aligned}$$



$= 0$

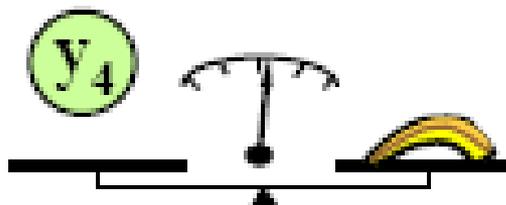
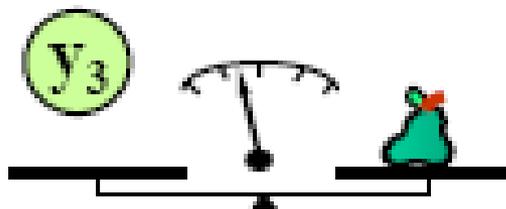
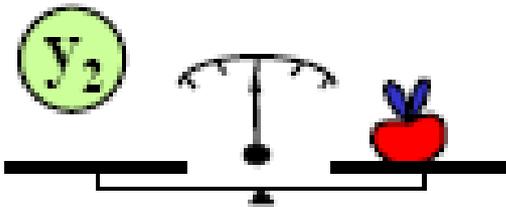
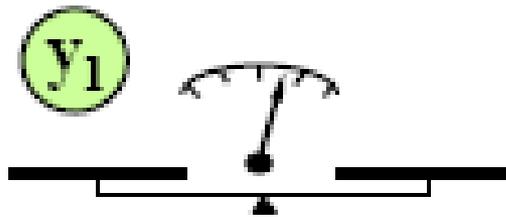
$= \sigma^2$
(variance de e_1)

$$\text{var}(\hat{m}) = 2\sigma^2$$

Pesée de plusieurs objets

- Si nous avons trois objets, comment les peser ?
- Quel est le prix minimum de l'expérimentation?

Expérimentateur n°1



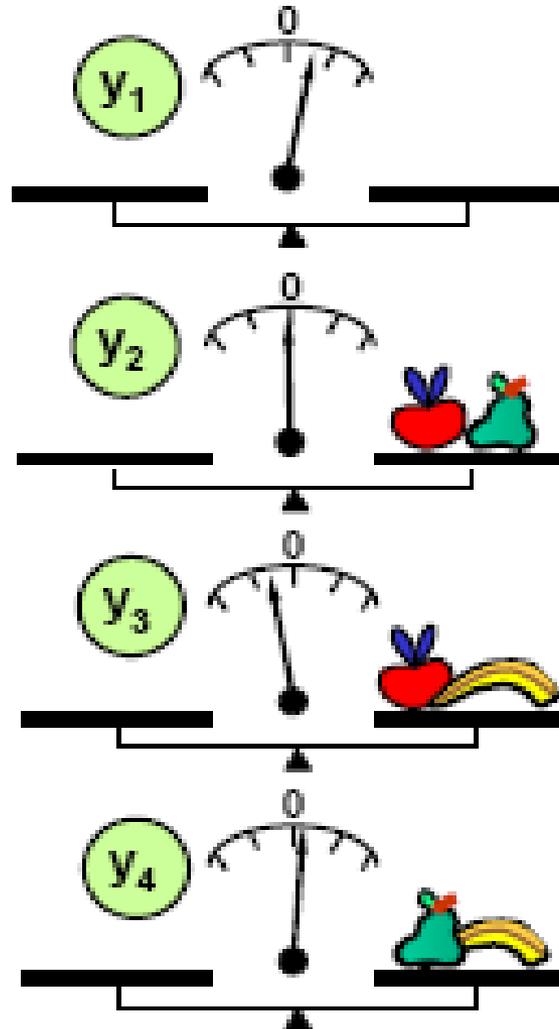
Lectures

Masse d'un objet : $\hat{m}_i = y_{i+1} - y_i$

$$\text{var}(\hat{m}_i) = \text{var}(Y_{i+1}) + \text{var}(Y_i)$$

$$\text{var}(\hat{m}_i) = 2\sigma^2$$

Expérimentateur n°2



Lectures

Masse d'un objet ?

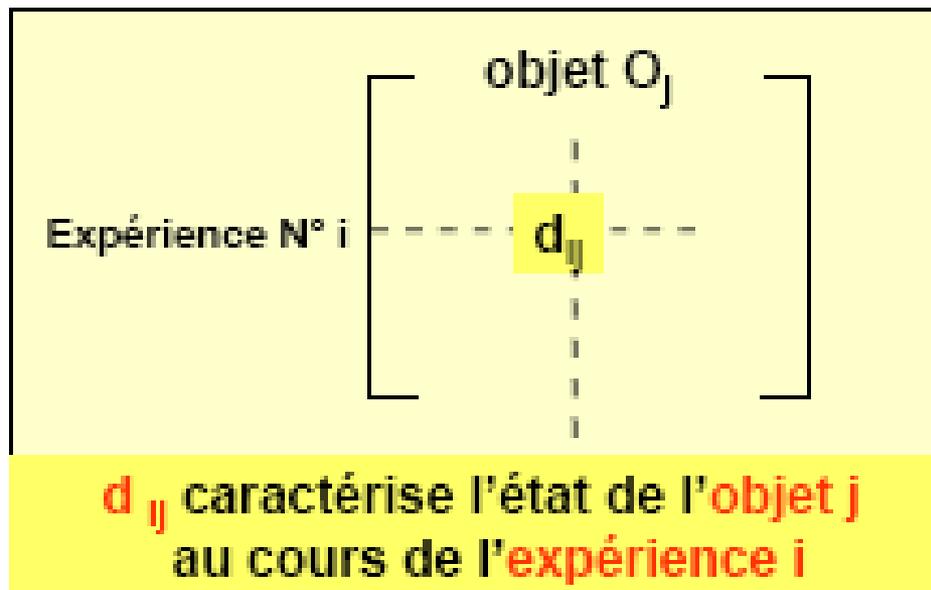
Difficile à lire directement



Introduction de la notion de matrice d'expériences

Matrice d'expériences

Une matrice d'expériences est un **tableau** permettant de **décrire une expérimentation** en donnant, pour chaque expérience, les valeurs des facteurs expérimentaux.



Expérimentateurs n°1 & n°2

Codage des objets : objet **absent** : état **0**
objet **présent** : état **1**

| Exp. | O ₁ | O ₂ | O ₃ |
|------|----------------|----------------|----------------|
| 1 | 0 | 0 | 0 |
| 2 | 1 | 0 | 0 |
| 3 | 0 | 1 | 0 |
| 4 | 0 | 0 | 1 |

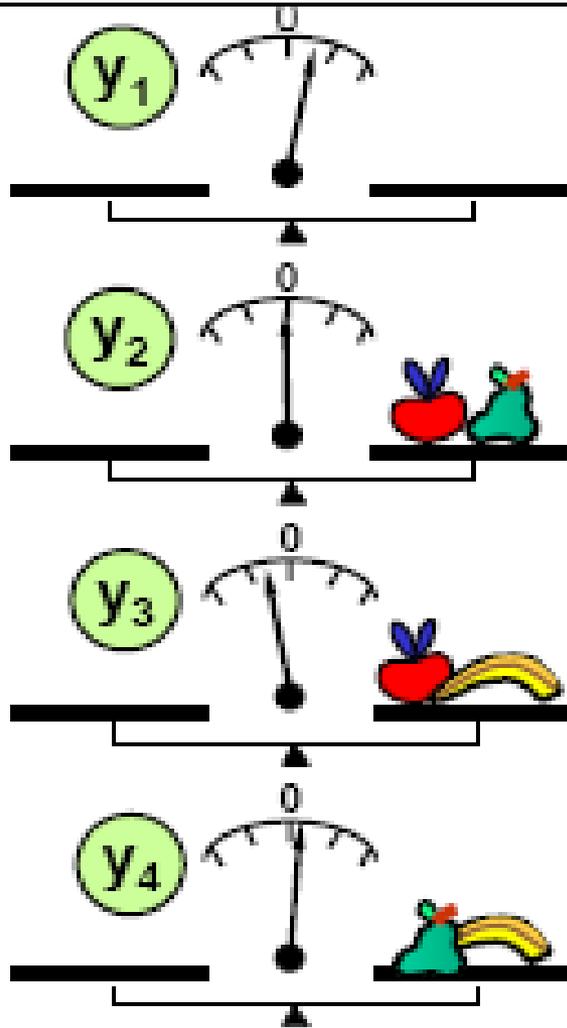
Plan N°1 (D₁)

| Exp. | O ₁ | O ₂ | O ₃ |
|------|----------------|----------------|----------------|
| 1 | 0 | 0 | 0 |
| 2 | 1 | 1 | 0 |
| 3 | 1 | 0 | 1 |
| 4 | 0 | 1 | 1 |

Plan N°2 (D₂)

Expérimentateur n° 2

Lectures



Systeme d'equations

$$\begin{aligned}y_1 &= \eta_0 + e_1 = \widehat{m}_0 \\y_2 &= \eta_1 + \eta_2 + e_2 = \widehat{m}_0 + \widehat{m}_1 + \widehat{m}_2 \\y_3 &= \eta_1 + \eta_3 + e_3 = \widehat{m}_0 + \widehat{m}_1 + \widehat{m}_3 \\y_4 &= \eta_2 + \eta_3 + e_4 = \widehat{m}_0 + \widehat{m}_2 + \widehat{m}_3\end{aligned}$$

Matrice d'expériences et système d'équations

| Exp. | O ₁ | O ₂ | O ₃ |
|------|----------------|----------------|----------------|
| 1 | 0 | 0 | 0 |
| 2 | 1 | 1 | 0 |
| 3 | 1 | 0 | 1 |
| 4 | 0 | 1 | 1 |

$$\begin{aligned}y_1 &= \widehat{m}_0 +0 \widehat{m}_1 +0 \widehat{m}_2 +0 \widehat{m}_3 \\y_2 &= \widehat{m}_0 +1 \widehat{m}_1 +1 \widehat{m}_2 +0 \widehat{m}_3 \\y_3 &= \widehat{m}_0 +1 \widehat{m}_1 +0 \widehat{m}_2 +1 \widehat{m}_3 \\y_4 &= \widehat{m}_0 +0 \widehat{m}_1 +1 \widehat{m}_2 +1 \widehat{m}_3\end{aligned}$$

Le **codage** utilisé pour écrire la matrice d'expériences permet de **faire correspondre**



les éléments de la matrice aux **coefficients des inconnues** du système d'équations



Inconnues = **masses à estimer**

Matrice du modèle et système d'équations

$$Y = m_0 + m_1 X_1 + m_2 X_2 + m_3 X_3$$

Avec la **notation matricielle**, on peut écrire de manière compacte le **système** de 4 équations linéaires qui lient les réponses observées aux états des objets :

$$Y = Xm$$

Vecteur réponse

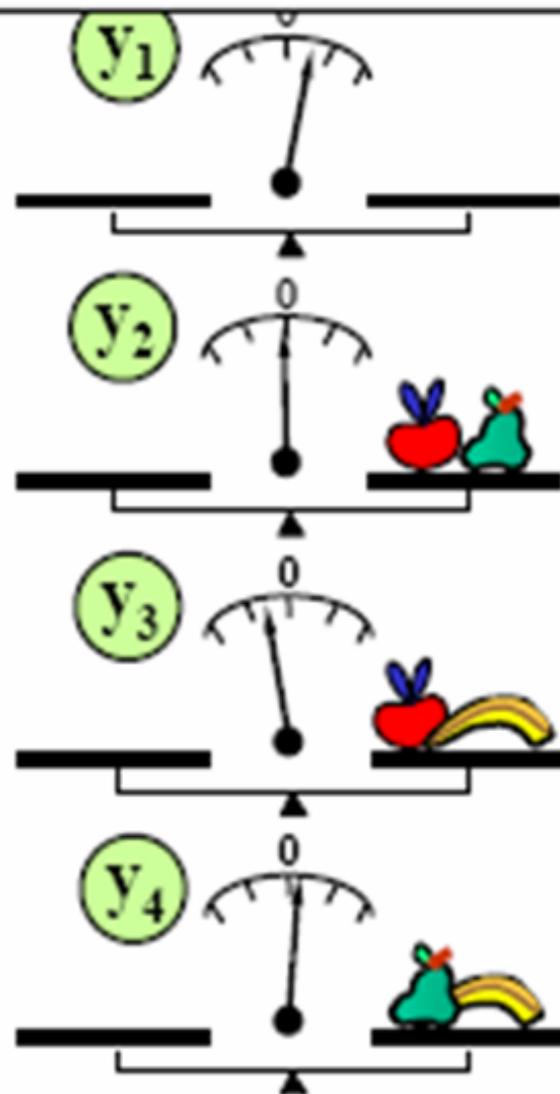
$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix} = \begin{bmatrix} 1 & -1 & -1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & 1 & 1 \end{bmatrix} * \begin{bmatrix} m_0 \\ m_1 \\ m_2 \\ m_3 \end{bmatrix}$$

Matrice du modèle

Vecteur des coefficients

Expérimentateur n°2

Lectures



Résolution

$$\hat{m}_0 = y_1$$

$$\hat{m}_1 = (-y_1 + y_2 + y_3 - y_4) / 2$$

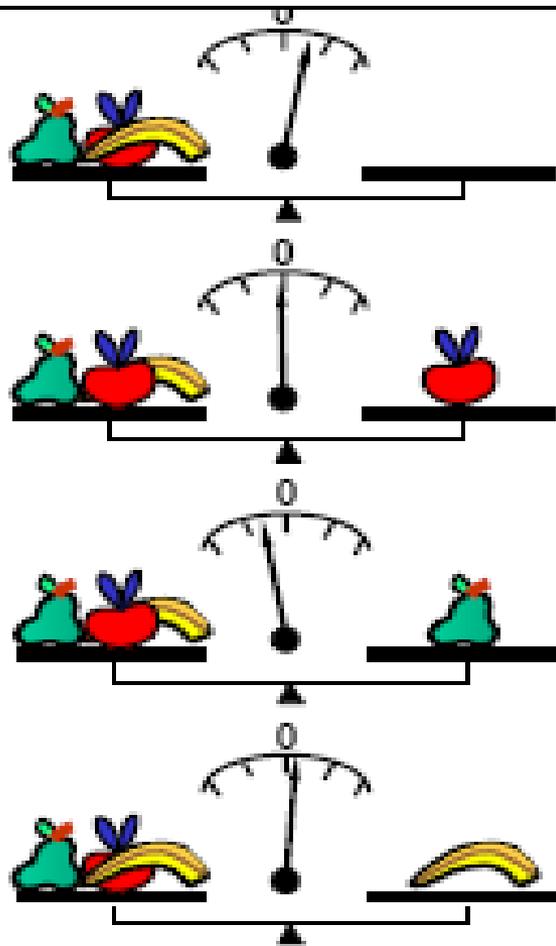
$$\hat{m}_2 = (-y_1 + y_2 - y_3 + y_4) / 2$$

$$\hat{m}_3 = (-y_1 - y_2 + y_3 + y_4) / 2$$

$$\begin{aligned} \text{var}(\hat{m}_i) &= \text{var}[1/2(\Sigma_{\text{algèbr. de 4 } Y_i})] \\ &= 1/4 [\text{var}(\Sigma_{\text{algèbr. de 4 } Y_i})] \end{aligned}$$

$$\text{var}(\hat{m}_i) = \sigma^2$$

Expérimentateur n°3



Codage des objets :

objet **absent** : état **0**

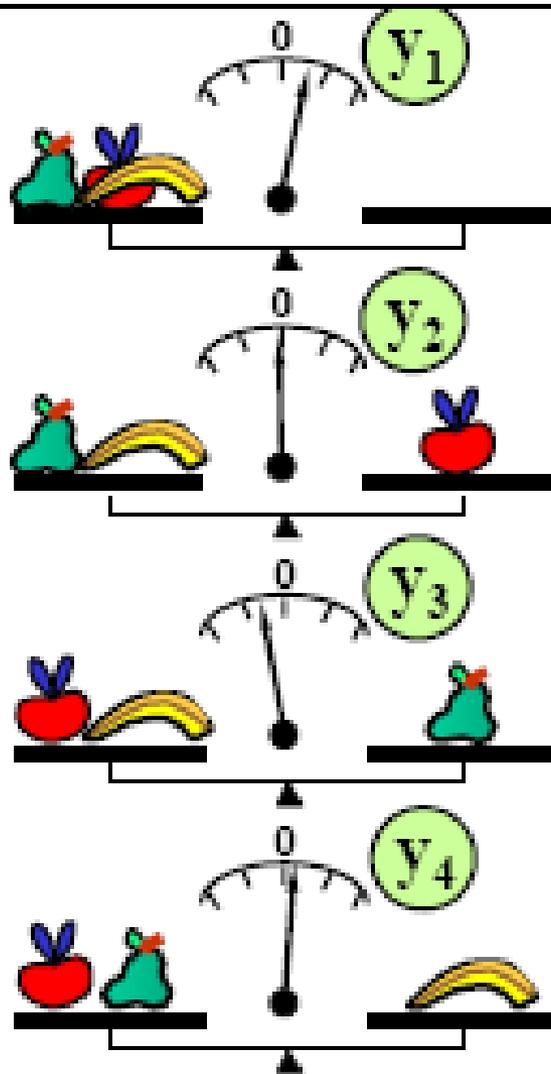
objet présent **à droite** : état **1**

objet présent **à gauche** : état **-1**

| Exp. | O ₁ | O ₂ | O ₃ |
|------|----------------|----------------|----------------|
| 1 | -1 | -1 | -1 |
| 2 | 1 | -1 | -1 |
| 3 | -1 | 1 | -1 |
| 4 | -1 | -1 | 1 |

Plan N°3 (D₃)

Expérimentateur n°3



Systeme d'equations

$$y_1 = \hat{m}_0 - \hat{m}_1 - \hat{m}_2 - \hat{m}_3$$

$$y_2 = \hat{m}_0 + \hat{m}_1 - \hat{m}_2 - \hat{m}_3$$

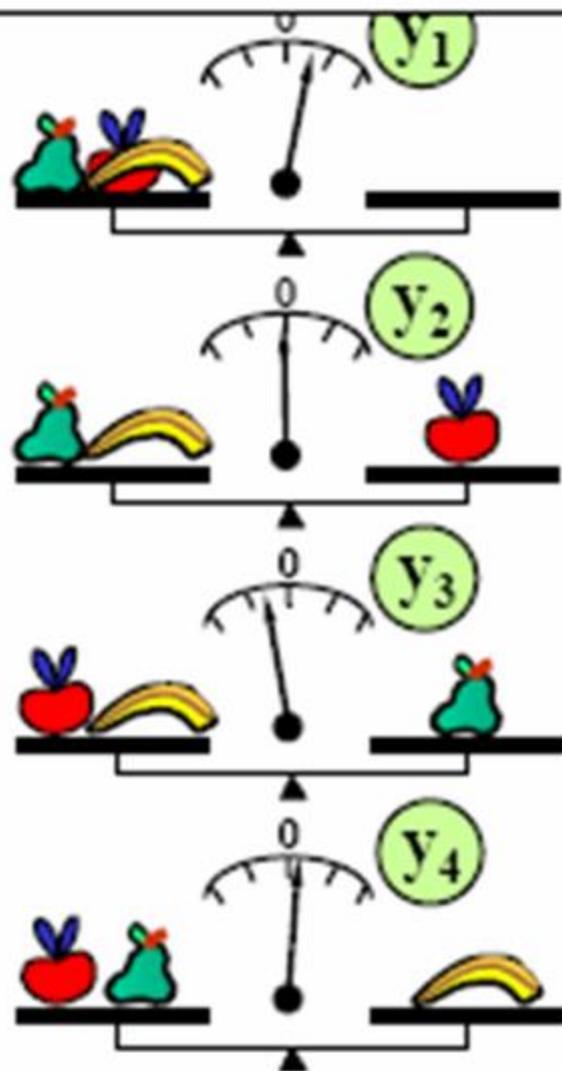
$$y_3 = \hat{m}_0 - \hat{m}_1 + \hat{m}_2 - \hat{m}_3$$

$$y_4 = \hat{m}_0 - \hat{m}_1 - \hat{m}_2 + \hat{m}_3$$

Lectures

Expérimentateur n°3

Lectures



Résolution

$$\hat{m}_0 = (-y_1 + y_2 + y_3 + y_4) / 2$$

$$\hat{m}_1 = (y_2 - y_1) / 2$$

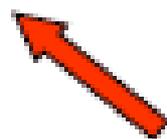
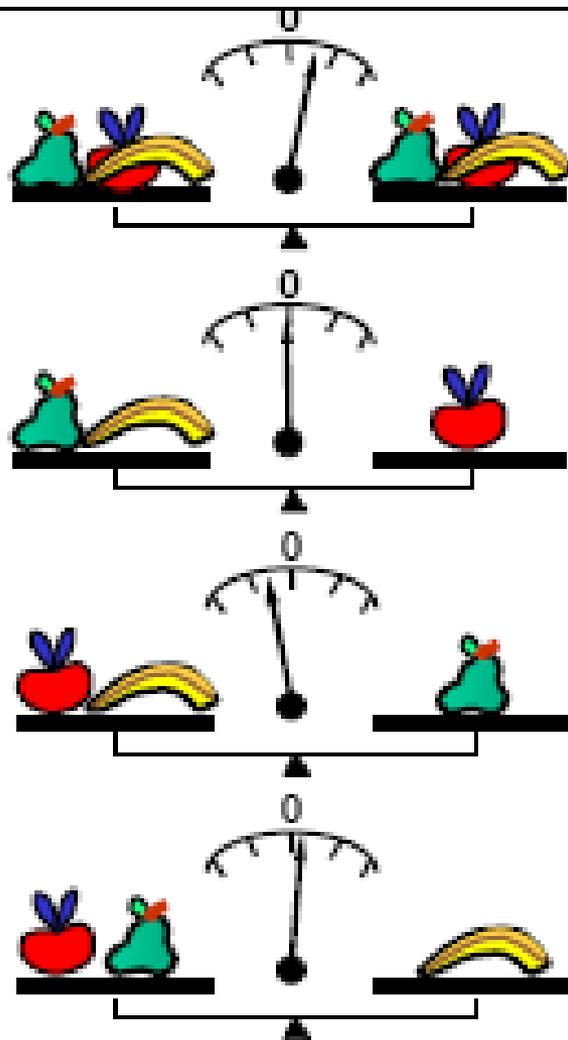
$$\hat{m}_2 = (y_3 - y_1) / 2$$

$$\hat{m}_3 = (y_4 - y_1) / 2$$

$$\begin{aligned} \text{var}(\hat{m}_i) &= \text{var}[1/2(\Sigma_{\text{algèbr. de 2 } Y_j})] \\ &= 1/4 [\text{var}(\Sigma_{\text{algèbr. de 2 } Y_j})] \end{aligned}$$

$$\text{var}(\hat{m}_i) = \sigma^2/2$$

Expérimentateur n°4



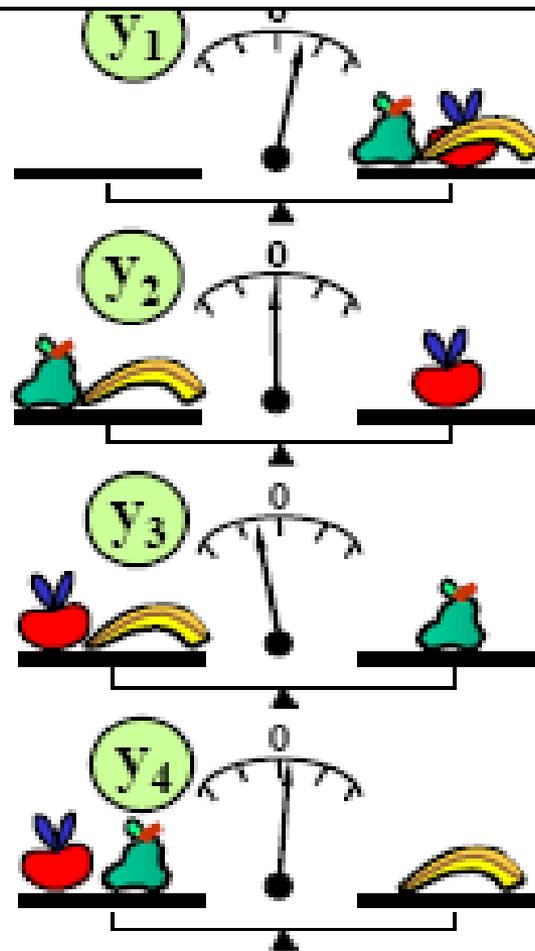
Seule l'expérience 1 change

| Exp. | O ₁ | O ₂ | O ₃ |
|------|----------------|----------------|----------------|
| 1 | 1 | 1 | 1 |
| 2 | 1 | -1 | -1 |
| 3 | -1 | 1 | -1 |
| 4 | -1 | -1 | 1 |

Plan N°4 (D₄)

Expérimentateur n°4

Lectures



Système d'équations

$$y_1 = \widehat{m}_0 + \widehat{m}_1 + \widehat{m}_2 + \widehat{m}_3$$

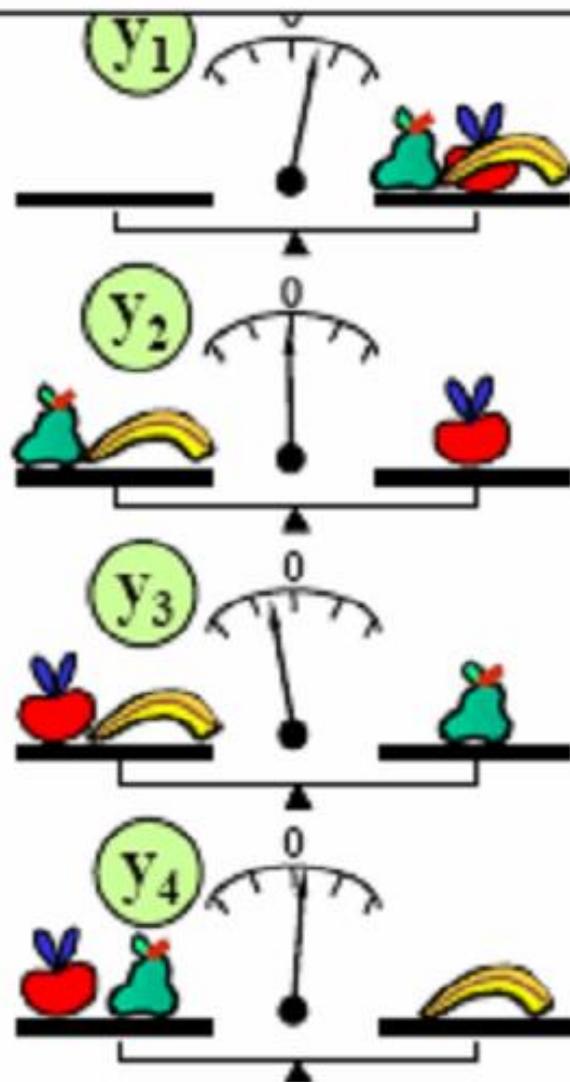
$$y_2 = \widehat{m}_0 + \widehat{m}_1 - \widehat{m}_2 - \widehat{m}_3$$

$$y_3 = \widehat{m}_0 - \widehat{m}_1 + \widehat{m}_2 - \widehat{m}_3$$

$$y_4 = \widehat{m}_0 - \widehat{m}_1 - \widehat{m}_2 + \widehat{m}_3$$

Expérimentateur n°4

Lectures



Résolution

$$\hat{m}_0 = (+y_1 + y_2 + y_3 + y_4) / 4$$

$$\hat{m}_1 = (+y_1 + y_2 - y_3 - y_4) / 4$$

$$\hat{m}_2 = (+y_1 - y_2 + y_3 - y_4) / 4$$

$$\hat{m}_3 = (+y_1 - y_2 - y_3 + y_4) / 4$$

$$\begin{aligned} \text{var}(\hat{m}_j) &= \text{var}[1/4(\Sigma_{\text{algèbr. de 4 } Y_i})] \\ &= 1/16 [\text{var}(\Sigma_{\text{algèbr. de 4 } Y_i})] \end{aligned}$$

$$\text{var}(\hat{m}_j) = \sigma^2/4$$

Conclusion

Nous n'avons **pas eu besoin des valeurs des y_i** pour prévoir la qualité des estimations.

La **qualité** de l'information expérimentale **ne dépend que du choix des essais** (de la matrice d'expériences).

Cette réflexion préalable peut être généralisée à toute expérimentation : **c'est avant d'expérimenter qu'il faut s'interroger sur la qualité** de l'expérimentation projetée.

THÉORÈME

Si on fait **N pesées**, la **variance** sur les coefficients est $\geq \sigma^2/N$

Comment obtenir une matrice optimale
pour minimiser la variance sur les masses estimées?

Modélisation

La relation entre la mesure Y (la lecture du cadran) et les différentes masses dont on cherche à déterminer la valeur est une relation linéaire qui s'écrit :

$$y_i = m_0 + m_1 X_{1i} + m_2 X_{2i} + m_3 X_{3i} + e_i$$

masses des objets

m_0, m_1, m_2 et m_3

état des masses dans
l'expérience j (la pesée)

X_1, X_2 et X_3 présence ou absence
sur les plateaux (-1, 1 ou 0).

erreur

Matrices d'Expériences

| Exp. | O ₁ | O ₂ | O ₃ |
|------|----------------|----------------|----------------|
| 1 | 0 | 0 | 0 |
| 2 | 1 | 0 | 0 |
| 3 | 0 | 1 | 0 |
| 4 | 0 | 0 | 1 |

Plan N°1 (D₁)

| Exp. | O ₁ | O ₂ | O ₃ |
|------|----------------|----------------|----------------|
| 1 | 0 | 0 | 0 |
| 2 | 1 | 1 | 0 |
| 3 | 1 | 0 | 1 |
| 4 | 0 | 1 | 1 |

Plan N°2 (D₂)

| Exp. | O ₁ | O ₂ | O ₃ |
|------|----------------|----------------|----------------|
| 1 | -1 | -1 | -1 |
| 2 | 1 | -1 | -1 |
| 3 | -1 | 1 | -1 |
| 4 | -1 | -1 | 1 |

Plan N°3 (D₃)

| Exp. | O ₁ | O ₂ | O ₃ |
|------|----------------|----------------|----------------|
| 1 | 1 | 1 | 1 |
| 2 | 1 | -1 | -1 |
| 3 | -1 | 1 | -1 |
| 4 | -1 | -1 | 1 |

Plan N°4 (D₄)

Expression matricielle de la régression

$$Y = m_0 + m_1 X_1 + m_2 X_2 + m_3 X_3 + e$$

Avec la **notation matricielle**, on peut écrire de manière compacte le **système** de 4 équations linéaires qui lient les réponses observées aux états des objets :

$$Y = X m + e$$

Vecteur réponse

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix} = \begin{bmatrix} 1 & -1 & -1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & 1 & 1 \end{bmatrix} * \begin{bmatrix} m_0 \\ m_1 \\ m_2 \\ m_3 \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \end{bmatrix}$$

Vecteur des erreurs

Matrice du modèle

Vecteur des coefficients

Expression matricielle de la régression

La matrice X représente la **matrice des coefficients des inconnues** m_j du système d'équations.

Calcul du vecteur des coefficients

X' étant la **matrice transposée** de X , si $X'X$ n'est pas singulière (c'est à dire s'il est possible de calculer la matrice inverse correspondante) :

on aura l'**estimation** du vecteur des coefficients du modèle au sens des moindres carrés par :

$$m = (X'X)^{-1}X'Y$$

Matrice de variance-covariance

Var (m) est une matrice où les **variances** des m_i sont disposées **sur la diagonale** et les **covariances de part et d'autre** de cette diagonale (matrice carrée symétrique) :

$$\text{var (m)} = \begin{bmatrix} \text{var (m}_0\text{)} & & & \\ \text{cov (m}_1\text{,m}_0\text{)} & \text{var (m}_1\text{)} & & \\ \text{cov (m}_2\text{,m}_0\text{)} & \text{cov (m}_2\text{,m}_1\text{)} & \text{var (m}_2\text{)} & \\ \text{cov (m}_3\text{,m}_0\text{)} & \text{cov (m}_3\text{,m}_1\text{)} & \text{cov (m}_3\text{,m}_2\text{)} & \text{var (m}_3\text{)} \end{bmatrix}$$
$$= \sigma^2 (\mathbf{X}'\mathbf{X})^{-1}$$

Qualité des Matrices d'Expériences

| Exp. | O ₁ | O ₂ | O ₃ |
|------|----------------|----------------|----------------|
| 1 | 0 | 0 | 0 |
| 2 | 1 | 0 | 0 |
| 3 | 0 | 1 | 0 |
| 4 | 0 | 0 | 1 |

Matrice D

| Exp. | O ₀ | O ₁ | O ₂ | O ₃ |
|------|----------------|----------------|----------------|----------------|
| 1 | 1 | 0 | 0 | 0 |
| 2 | 1 | 1 | 0 | 0 |
| 3 | 1 | 0 | 1 | 0 |
| 4 | 1 | 0 | 0 | 1 |

Matrice X du modèle :

$$Y = m_0 + \sum m_i X_i$$

masse de l'objet i

état de l'objet i
dans la pesée considérée

Matrice de **variance-covariance** : $(X'X)^{-1}$

Matrice de variance-covariance

| Exp. | O ₀ | O ₁ | O ₂ | O ₃ |
|------|----------------|----------------|----------------|----------------|
| 1 | 1 | 0 | 0 | 0 |
| 2 | 1 | 1 | 0 | 0 |
| 3 | 1 | 0 | 1 | 0 |
| 4 | 1 | 0 | 0 | 1 |

Plan N°1

Matrice X

| | m ₀ | m ₁ | m ₂ | m ₃ |
|----------------|----------------|----------------|----------------|----------------|
| m ₀ | 1 | -1 | -1 | -1 |
| m ₁ | -1 | 2 | 1 | 1 |
| m ₂ | -1 | 1 | 2 | 1 |
| m ₃ | -1 | 1 | 1 | 2 |

Matrice (X'X)⁻¹[var.-covar.]

| Exp. | O ₀ | O ₁ | O ₂ | O ₃ |
|------|----------------|----------------|----------------|----------------|
| 1 | 1 | 0 | 0 | 0 |
| 2 | 1 | 1 | 1 | 0 |
| 3 | 1 | 1 | 0 | 1 |
| 4 | 1 | 0 | 1 | 1 |

Plan N°2

Matrice X

| | m ₀ | m ₁ | m ₂ | m ₃ |
|----------------|----------------|----------------|----------------|----------------|
| m ₀ | 1 | -0.5 | -0.5 | -0.5 |
| m ₁ | -0.5 | 1 | 0 | 0 |
| m ₂ | -0.5 | 0 | 1 | 0 |
| m ₃ | -0.5 | 0 | 0 | 1 |

Matrice (X'X)⁻¹[var.-covar.]

Matrice de variance-covariance

| Exp. | O ₀ | O ₁ | O ₂ | O ₃ |
|------|----------------|----------------|----------------|----------------|
| 1 | 1 | -1 | -1 | -1 |
| 2 | 1 | 1 | -1 | -1 |
| 3 | 1 | -1 | 1 | -1 |
| 4 | 1 | -1 | -1 | 1 |

Plan N°3

Matrice X

| | m ₀ | m ₁ | m ₂ | m ₃ |
|----------------|----------------|----------------|----------------|----------------|
| m ₀ | 1 | -0,5 | -0,5 | -0,5 |
| m ₁ | -0,5 | 0,5 | 0,25 | 0,25 |
| m ₂ | -0,5 | 0,25 | 0,5 | 0,25 |
| m ₃ | -0,5 | 0,25 | 0,25 | 0,5 |

Matrice (X'X)⁻¹[var.-covar.]

| Exp. | O ₀ | O ₁ | O ₂ | O ₃ |
|------|----------------|----------------|----------------|----------------|
| 1 | 1 | 1 | 1 | 1 |
| 2 | 1 | 1 | -1 | -1 |
| 3 | 1 | -1 | 1 | -1 |
| 4 | 1 | -1 | -1 | 1 |

Plan N°4

Matrice X

| | m ₀ | m ₁ | m ₂ | m ₃ |
|----------------|----------------|----------------|----------------|----------------|
| m ₀ | 0,25 | 0 | 0 | 0 |
| m ₁ | 0 | 0,25 | 0 | 0 |
| m ₂ | 0 | 0 | 0,25 | 0 |
| m ₃ | 0 | 0 | 0 | 0,25 |

Matrice (X'X)⁻¹[var.-covar.]

Calcul des coefficients : Régression

Statistiques de la Régression

| | |
|---|--------|
| Coefficient de détermination multiple | 0,9244 |
| Coefficient de détermination R^2 | 0,8545 |
| Coefficient de détermination ajusté R^2_a | 0,6604 |
| Erreur-type | 6,1364 |
| Observations | 8 |

Analyse Statistique des résultats

Analyse de la variance

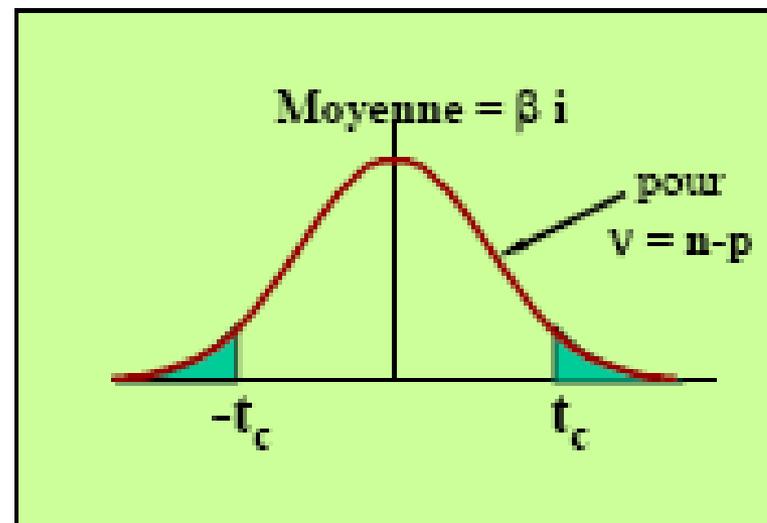
| Source de la Variation | Sommes des carrés | Degrés de liberté | Carrés moyens | F observé |
|------------------------|-------------------|-------------------|------------------------|--------------|
| Régression (liaison) | SCE_L | $p - 1$ | $CM_L = SCE_L / (p-1)$ | CM_L / s^2 |
| Résidus | SCE_R | $n - p$ | $CM_R = SCE_R / (n-1)$ | |
| Totale | SCE_T | $n-1$ | | |

| Sources de variation | Sommes des Carrés | Degr. de lib. | Carrés moyens | $F_{obs.}$ | Signif. |
|----------------------|-------------------|---------------|---------------|-----------------------|---------|
| Régression | 663,20 | $5-1=4$ | 165,80 | 4,403 | 0,127 |
| Résidus | 112,97 | $8-5=3$ | 37,66 | $F_{4;3;0,05} = 9,12$ | |
| Total | 776,16 | $8-1=7$ | | | |

Analyse Statistique des résultats

Analyse des coefficients

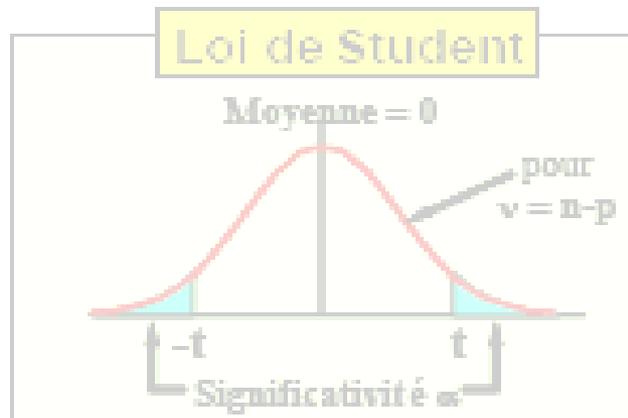
Le coefficient b_i est distribué selon une **distribution de Student** de **moyenne β_i** , d'**écart-type $e.t.(b_i)$** et $(n-p)$ degrés de liberté.



Intervalle de confiance pour b_i :

$$b_i \pm t_c e.t.(b_i)$$

Significativité des coefficients



$$t = \frac{b_i - \beta_i^0}{\text{é.t.}(b_i)}$$

$$t = \frac{b_i}{\text{é.t.}(b_i)}$$

Analyse des coefficients

| | Coefficients | Erreur | Statist. t | Probab. (signif.) | Seuil de conf. 95% | |
|-----------|--------------|--------|------------|----------------------|--------------------|----------|
| | | | | | Lim.inf. | Lim.sup. |
| Constante | 52,45 | 9,15 | 5,73 | 0,011 | 23,34 | 81,56 |
| M_1 | -0,38 | 0,22 | -1,74 | 0,180 | -1,07 | 0,31 |
| M_2 | -0,44 | 0,43 | -1,03 | 0,381 | -1,83 | 0,94 |
| S_1 | -1,02 | 0,29 | -3,53 | 0,039 | -1,94 | -0,10 |
| S_2 | -0,45 | 0,43 | -1,05 | 0,371 | -1,84 | 0,93 |

Calcul des effets selon Pareto

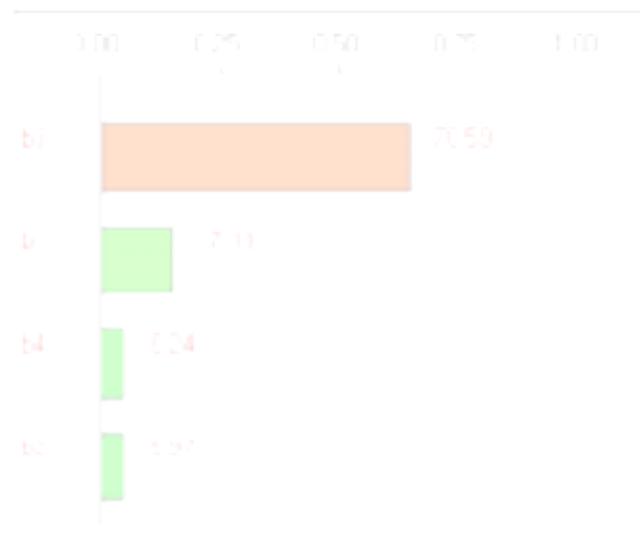
Il s'agit d'évaluer la participation de chaque effet du modèle par rapport à l'ensemble des effets.

Le pourcentage de l'effet selon Pareto est obtenu en élevant le coefficient au carré et en divisant par la somme des carrés des coefficients :

| Coefficients | Carrés des coefficients | Effets Pareto |
|------------------|-------------------------|---------------|
| b_1 -3.775 | 14.25 | 17.19 |
| b_2 -2.225 | 4.95 | 5.97 |
| b_3 -7.65 | 58.52 | 70.59 |
| b_4 -2.275 | 5.18 | 6.24 |
| somme des carrés | | |
| 82.90 | | |

Calcul des effets selon Pareto

Effets selon Pareto



Effets Pareto cumulés





Chapitre III

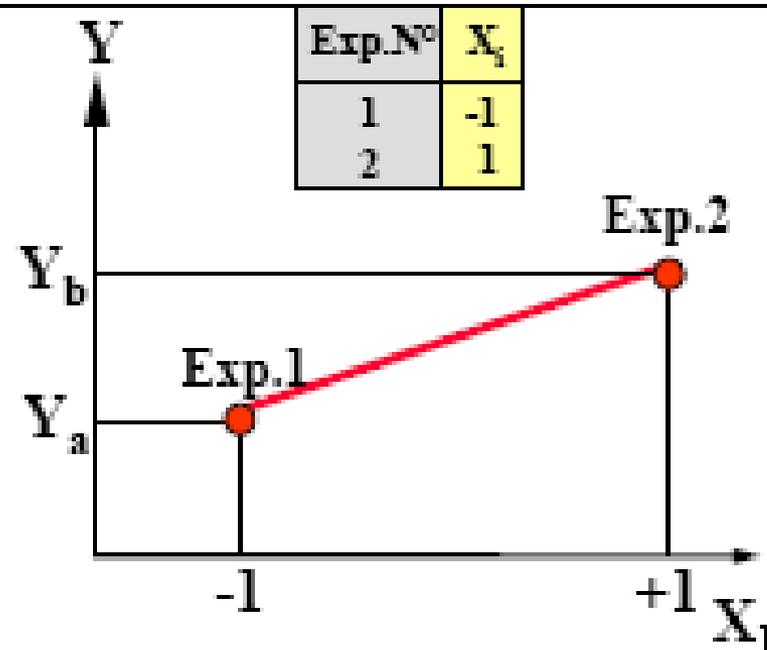
Matrice d'Hadamard
Le criblage des facteurs :
classement hiérarchisé des
facteurs

Effets des Facteurs

**Effet
d'un facteur**

Effet de X_1 : E_1

$$E_1 = Y_b - Y_a$$



Effet de X :

(valeur de la réponse Y au niveau +1)

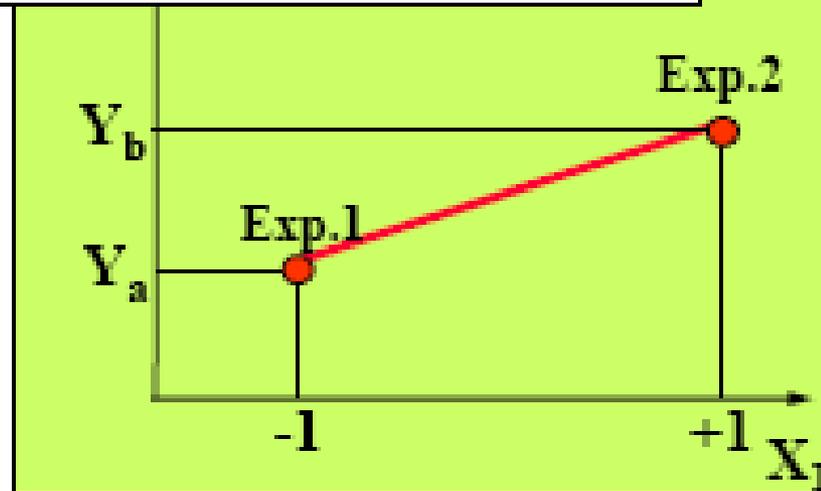
-

(Valeur de la réponse Y au niveau -1)

Poids des Facteurs

**Poids
d'un facteur
(pente)**

| Exp.N° | X_1 |
|--------|-------|
| 1 | -1 |
| 2 | 1 |



$$\text{Modèle : } Y = b_0 + b_1 X_1$$

$$b_1 = \frac{Y_b - Y_a}{2}$$

« poids du facteur X_1 » $b_1 = \frac{E_1}{2}$

$$b_0 = \frac{Y_b + Y_a}{2}$$

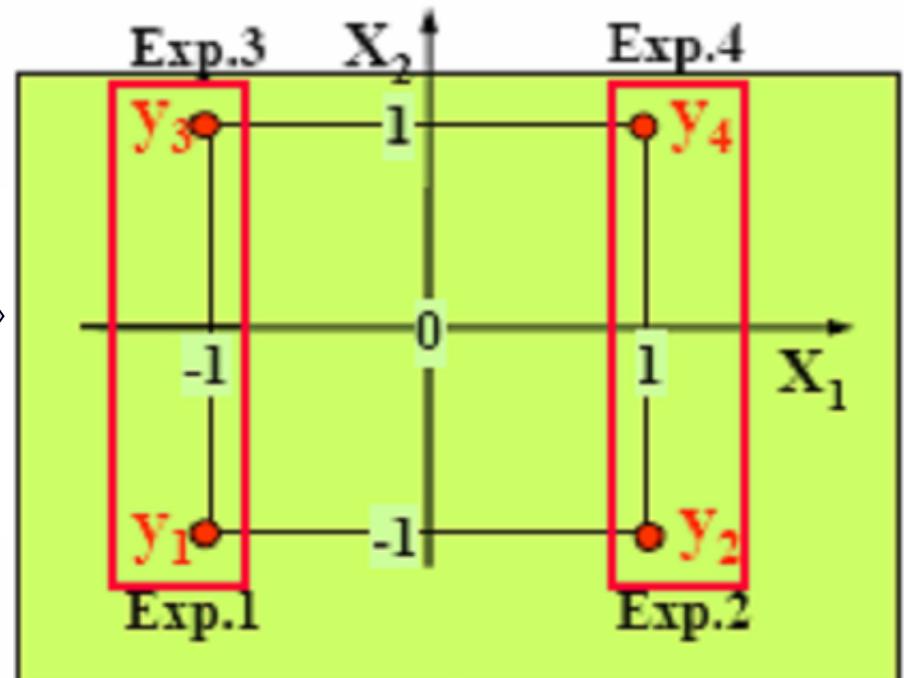
« valeur moyenne » "réponse théorique"
au centre du domaine

Effets des Facteurs

Cas de deux facteurs

| Exp.N° | X ₁ | X ₂ |
|--------|----------------|----------------|
| 1 | -1 | -1 |
| 2 | 1 | -1 |
| 3 | -1 | 1 |
| 4 | 1 | 1 |

Pour X₁



Réponse moyenne pour X₁ = -1
 $(y_1 + y_3) / 2$

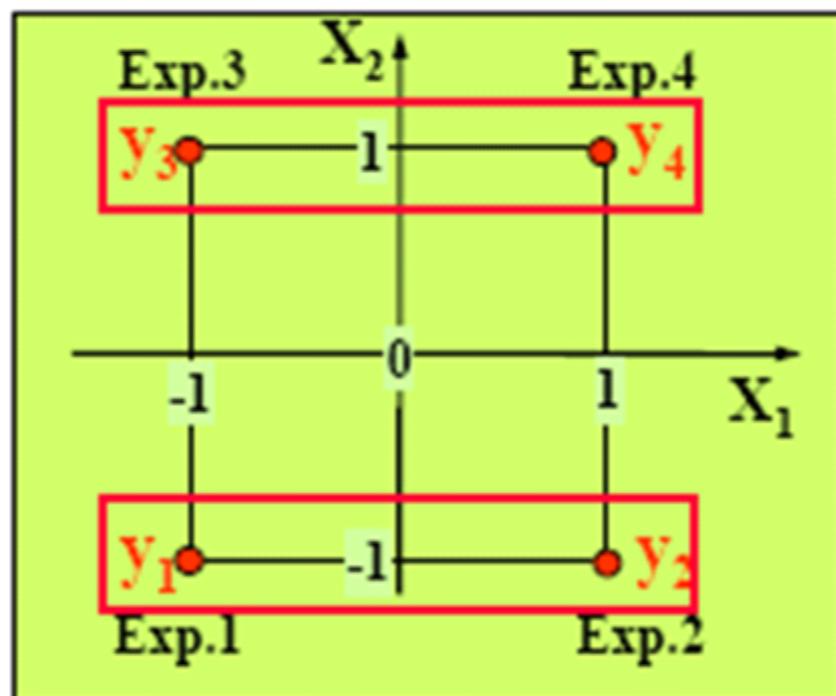
Réponse moyenne pour X₁ = 1
 $(y_2 + y_4) / 2$

$E_1/2 =$ Poids du facteur X₁ :
 $[(y_2 + y_4) / 2] - [(y_1 + y_3) / 2]$



poids de X₁ : $E_1/2 = \frac{-y_1 + y_2 - y_3 + y_4}{2}$

Pour le cas de X_2



$E_2 / 2 =$ Poids du facteur X_2 :
 $[(y_3+y_4)/2] - [(y_1+y_2)/2]$

Réponse moyenne
pour $X_2 = -1$
 $(y_1+y_2)/2$

Réponse moyenne
pour $X_2 = 1$
 $(y_3+y_4)/2$

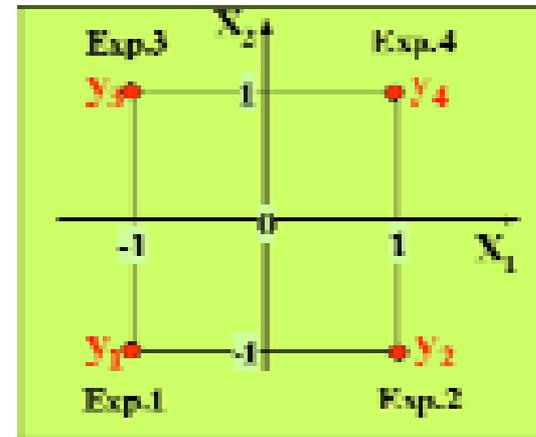
poids de X_2 :

$$E_2 / 2 = \frac{-y_1 - y_2 + y_3 + y_4}{2}$$

Poids des Facteurs

**Poids pour
deux facteurs**

| Exp.N° | X ₁ | X ₂ |
|--------|----------------|----------------|
| 1 | -1 | -1 |
| 2 | 1 | -1 |
| 3 | -1 | 1 |
| 4 | 1 | 1 |



Modèle : $Y = b_0 + b_1 X_1 + b_2 X_2$

Poids de X₁ $b_1 = \frac{E_1}{2} = \frac{-Y_1 + Y_2 - Y_3 + Y_4}{4}$

Poids de X₂ $b_2 = \frac{E_2}{2} = \frac{-Y_1 - Y_2 + Y_3 + Y_4}{4}$

Valeur moyenne $b_0 = \frac{+Y_1 + Y_2 + Y_3 + Y_4}{4}$

Poids = Pente = Coefficient

Modèle pour **k facteurs** :

$$Y = b_0 + b_1 X_1 + b_2 X_2 + b_3 X_3 + \dots + b_k X_k$$

Rappel du théorème des plans de pesée :

Si on fait **N pesées**, la **variance** sur les coefficients est $\geq \sigma^2/N$

Pour avoir la **variance des estimations minimale** σ^2/N , la **condition nécessaire et suffisante** est que l'on ait :

$$X'X = N I_N$$

Plans de Pesées et Matrices d 'Hadamard

Cette propriété est obtenue avec les **matrices d'Hadamard**, qui ne contiennent que des valeurs -1 ou +1 .

Elles n'existent que pour **N multiple de 4** :
4, 8, 12, 16, 20 ...



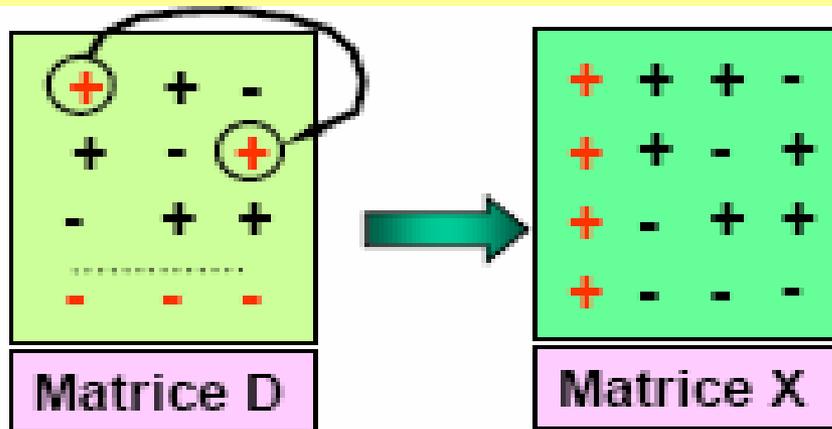
Jacques Salomon
HADAMARD
(1865 - 1963)

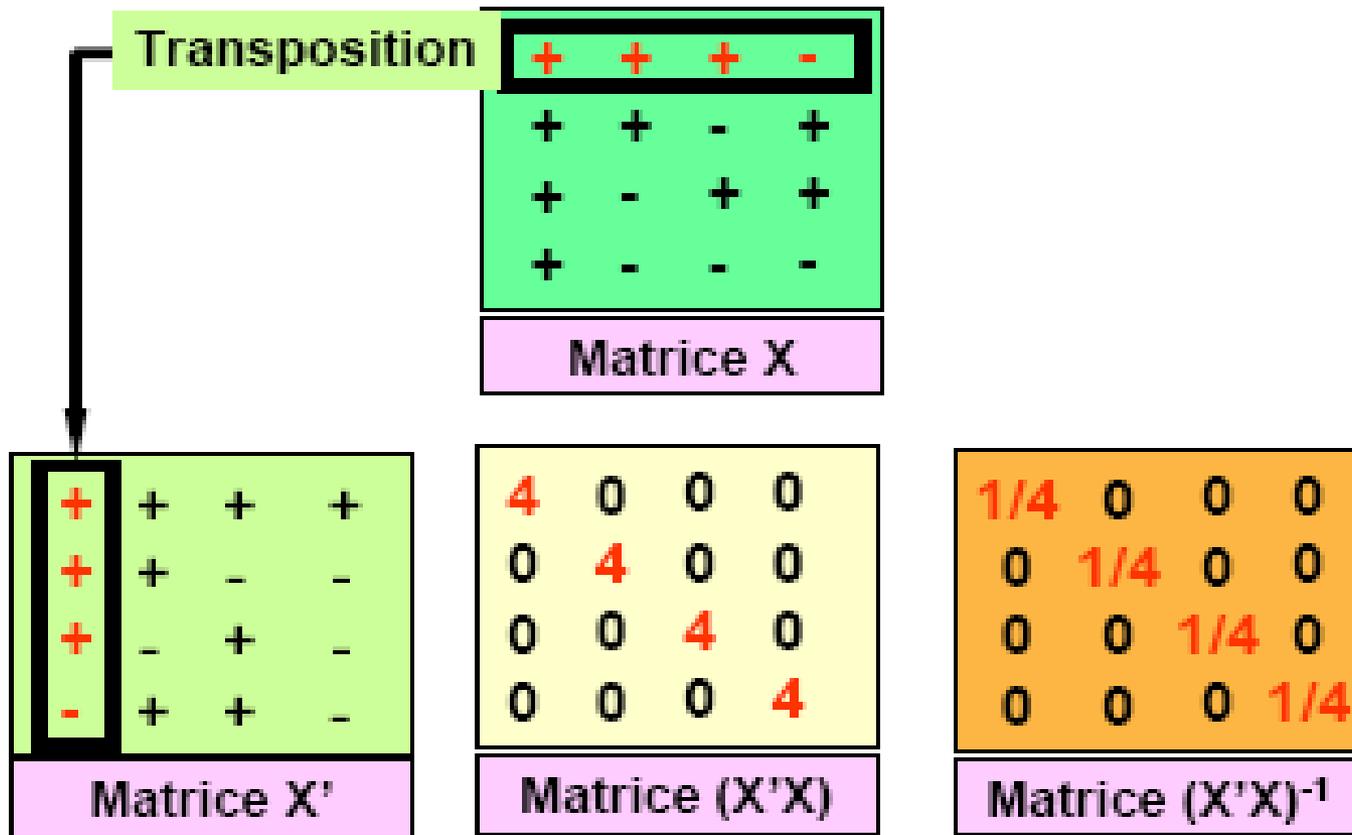
k : Nombre de Facteurs
 N : Nombre de pesées

Construction de Plackett et Burman

- k ≤ 3 (N=4)** : + + -
- k ≤ 7 (N=8)** : + + + - + - -
- k ≤ 11 (N=12)** : + + - + + + - - - + -
- k ≤ 15 (N=16)** : + + + + - + - + + - - + - - -
- k ≤ 19 (N=20)** : + + - - + + + + - + - + - - - - + + -
- k ≤ 23 (N=24)** : + + + + + - + - + + - - + + - - + - + - - - -

Exemple N = 4
 après N-2 permutations





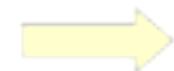
A



| | | | |
|---|---|---|---|
| + | + | + | - |
| + | + | - | + |
| + | - | + | + |
| + | - | - | - |

Matrice X

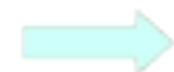
C



| | | | |
|---|---|---|---|
| + | - | - | + |
| + | - | + | - |
| + | + | - | - |
| + | + | + | + |

Matrice X

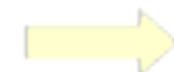
B



| | | | |
|---|---|---|---|
| + | + | + | - |
| + | - | + | + |
| + | + | - | + |
| + | - | - | - |

Matrice X

D



| | | | |
|---|---|---|---|
| + | - | - | + |
| + | + | - | - |
| + | - | + | - |
| + | + | + | + |

Matrice X

Plan de pesée : exemple

Digestion enzymatique on-line

Sept variables opératoires sont évaluées simultanément

- Débit
- Température
- Quantité injectée
- Force ionique du tampon
- Traitement préalable de l'échantillon
- Présence de SDS dans l'échantillon
- Type de tampon

Variations des conditions expérimentales

| facteur | Débit (ml/mn) | Temp. (°C) | Qté inj. (mg) | Conc. (mM) | TTT | SDS (%) | Tamp. |
|----------|------------------|---------------|------------------|---------------|-----|------------|--------|
| minimum | 0.5 | 30° | 0.5 | 20 | OUI | Sans | Phosp. |
| maximum | 3.0 | 37° | 1.5 | 100 | NON | 0.05 | Ambic |
| Standard | 1.0 | 30° | 1.0 | 60 | OUI | 0.05 | Phosp. |

Facteurs

X_1/U_1 : Débit (ml/min)

X_2/U_2 : Température (°C)

X_3/U_3 : Qté injectée (mg)

X_4/U_4 : Concentration tampon (mM)

X_5/U_5 : Traitement (TTT)

X_6/U_6 : % SDS

X_7/U_7 : Nature du tampon

Minimum (-)

Niveau -1

0.5

30

0.5

20

Oui

Sans

Phosphate

Maximum (+)

Niveau 1

3.0

37

1.5

60

Non

0.05

Ambic

Matrice d'expérience

$k \leq 7$ (N=8)

: + + + - + - -

Décalage droite

| N° Exp | X1 | X2 | X3 | X4 | X5 | X6 | X7 |
|--------|----|----|----|----|----|----|----|
| 1 | 1 | 1 | 1 | -1 | 1 | -1 | -1 |
| 2 | | 1 | 1 | 1 | -1 | 1 | -1 |
| 3 | | | | | | | |
| 4 | | | | | | | |
| 5 | | | | | | | |
| 6 | | | | | | | |
| 7 | | | | | | | |
| 8 | | | | | | | |

-1

Matrice d'expérience

$k \leq 7$ (N=8)

: + + + - + - -

Décalage droite

| N° Exp | X1 | X2 | X3 | X4 | X5 | X6 | X7 |
|--------|----|----|----|----|----|----|----|
| 1 | 1 | 1 | 1 | -1 | 1 | -1 | -1 |
| 2 | -1 | 1 | 1 | 1 | -1 | 1 | -1 |
| 3 | | -1 | 1 | 1 | 1 | -1 | 1 |
| 4 | | | | | | | |
| 5 | | | | | | | |
| 6 | | | | | | | |
| 7 | | | | | | | |
| 8 | | | | | | | |

-1

Matrice d'expérience

$k \leq 7$ (N=8)

: + + + - + - -

| N° Exp | X1 | X2 | X3 | X4 | X5 | X6 | X7 |
|--------|----|----|----|----|----|----|----|
| 1 | 1 | 1 | 1 | -1 | 1 | -1 | -1 |
| 2 | -1 | 1 | 1 | 1 | -1 | 1 | -1 |
| 3 | -1 | -1 | 1 | 1 | 1 | -1 | 1 |
| 4 | 1 | -1 | -1 | 1 | 1 | 1 | -1 |
| 5 | -1 | 1 | -1 | -1 | 1 | 1 | 1 |
| 6 | 1 | -1 | 1 | -1 | -1 | 1 | 1 |
| 7 | 1 | 1 | -1 | 1 | -1 | -1 | 1 |
| 8 | -1 | -1 | -1 | -1 | -1 | -1 | -1 |

Dernière ligne : - - - - -

Matrice complète

Plan d'expérimentation

| N° | X1 | X2 | X3 | X4 | X5 | X6 | X7 |
|----|----|----|----|----|----|----|----|
| 1 | 1 | 1 | 1 | -1 | 1 | -1 | -1 |
| 2 | -1 | 1 | 1 | 1 | -1 | 1 | -1 |
| 3 | -1 | -1 | 1 | 1 | 1 | -1 | 1 |
| 4 | 1 | -1 | -1 | 1 | 1 | 1 | -1 |
| 5 | -1 | 1 | -1 | -1 | 1 | 1 | 1 |
| 6 | 1 | -1 | 1 | -1 | -1 | 1 | 1 |
| 7 | 1 | 1 | -1 | 1 | -1 | -1 | 1 |
| 8 | -1 | -1 | -1 | -1 | -1 | -1 | -1 |

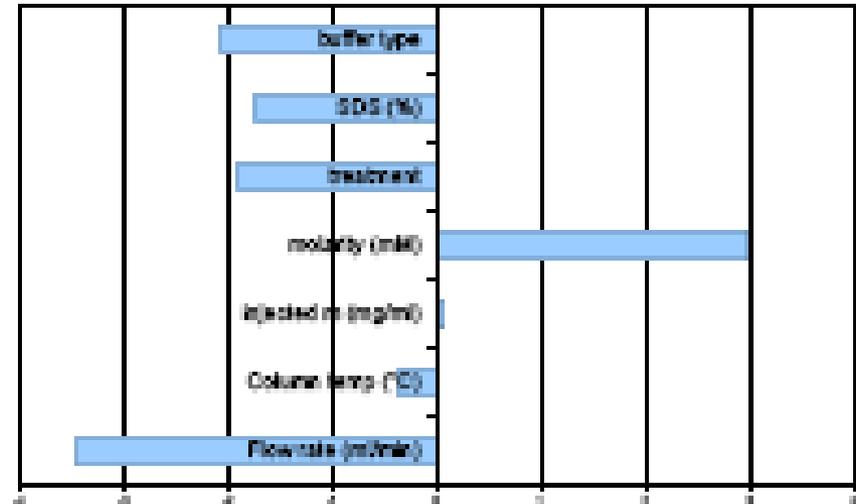
| N° | U1 | U2 | U3 | U4 | U5 | U6 | U7 |
|----|-----|----|-----|-----|-----|------|----|
| 1 | 3 | 37 | 1.5 | 20 | non | 0 | P |
| 2 | 0.5 | 37 | 1.5 | 100 | oui | 0.05 | P |
| 3 | 0.5 | 30 | 1.5 | 100 | non | 0 | A |
| 4 | 3 | 30 | 0.5 | 100 | non | 0.05 | P |
| 5 | 0.5 | 37 | 0.5 | 20 | non | 0.05 | A |
| 6 | 3 | 30 | 1.5 | 20 | oui | 0.05 | A |
| 7 | 3 | 37 | 0.5 | 100 | oui | 0 | A |
| 8 | 0.5 | 30 | 0.5 | 20 | oui | 0 | P |

| Facteurs | Niveau -1(-) | Niveau 1(+) |
|---------------------------------------|---------------|-------------|
| X_1/U_1 : Débit (ml/min) | 0.5 | 3.0 |
| X_2/U_2 : Température (°C) | 30 | 37 |
| X_3/U_3 : Qté Injectée (mg) | 0.5 | 1.5 |
| X_4/U_4 : Concentration tampon (mM) | 20 | 60 |
| X_5/U_5 : Traitement (TTT) | Oui | Non |
| X_6/U_6 : % SDS | Sans | 0.05 |
| X_7/U_7 : Nature du tampon | Phosphate (P) | Ambic (A) |

Coefficients

Coefficients

| | |
|--------------------|----------|
| Constante | 11.97125 |
| Flow rate (ml/min) | -3.44875 |
| Column temp (°C) | -0.34625 |
| injected m (mg/ml) | 0.07125 |
| molarity (mM) | 2.98375 |
| treatment | -1.89125 |
| SDS (%) | -1.73375 |
| buffer type | -2.07375 |



Significativité des coefficients :

Estimation de la variance expérimentale par l'utilisation de la répétition dans les conditions standards

Significativité des coefficients

Répétitions : estimations de la variance expérimentale
(pas de point central : valeurs qualitatives)

| | |
|----------|-------|
| PC1 | 11.4 |
| PC2 | 13.4 |
| PC3 | 9.2 |
| PC4 | 10.8 |
| PC5 | 13.5 |
| PC6 | 8.5 |
| Variance | 4.3 |
| CV | 18.4% |

$$\text{é.t.}(b_i) = \frac{S}{\sqrt{n}}$$

| | |
|--------------------|-------|
| Variance exp | 4.33 |
| ecartype exp. | 2.08 |
| <u>Erreur-type</u> | 0.736 |

La **significativité** va être **déterminée en prenant** :

$$t = \frac{b_i}{\text{é.type}(b_i)}$$

Significativité des coefficients

| | Coefficient | Erreur-type | Statistique t | Probabilité |
|--------------------|-------------|-------------|---------------|--------------------------|
| Constante | 11.87125 | 0.736 | 16.272 | 0.0% |
| Flow rate (ml/min) | -5.44875 | 0.736 | 4.698 | 0.5% très différent de 0 |
| Column temp (°C) | -0.54825 | 0.736 | 0.471 | 66.8% = 0 |
| injected m (µg/ml) | 0.07125 | 0.736 | 0.097 | 92.7% = 0 |
| molarity (µM) | 2.96375 | 0.736 | 4.050 | 1.0% très différent de 0 |
| treatment | -1.89125 | 0.736 | 2.571 | 5.0% A voir |
| SDS (%) | -1.78875 | 0.736 | 2.357 | 6.5% A voir |
| buffer type | -2.07375 | 0.736 | 2.819 | 3.7% A voir |

Les effets importants sont :

1. Le **débit** de la phase mobile
2. Le **type** et la **molarité** du **tampon**

Les effets possibles sont:

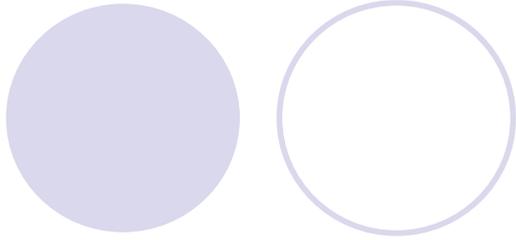
1. Le **traitement** de l'échantillon
2. La présence de **SDS**

Ne présentent pas d'effet

1. La **température** de la colonne
2. La **masse** injectée

Plus la probabilité est élevée, moins le coefficient est significatif

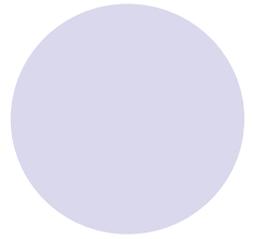
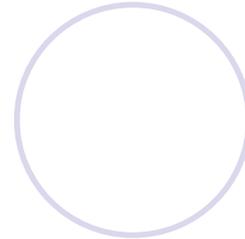
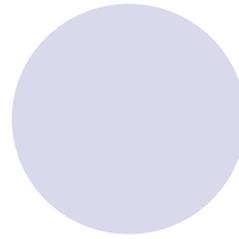
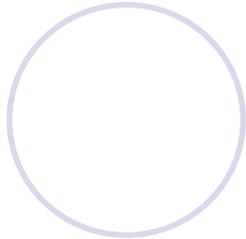
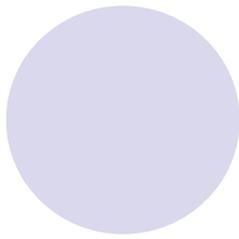
Valeur seuil : 5 %



Chapitre IV



Plans factoriels complets **FFD**



Matrices : pour le cas de deux facteurs

Effet d'interaction

synergique

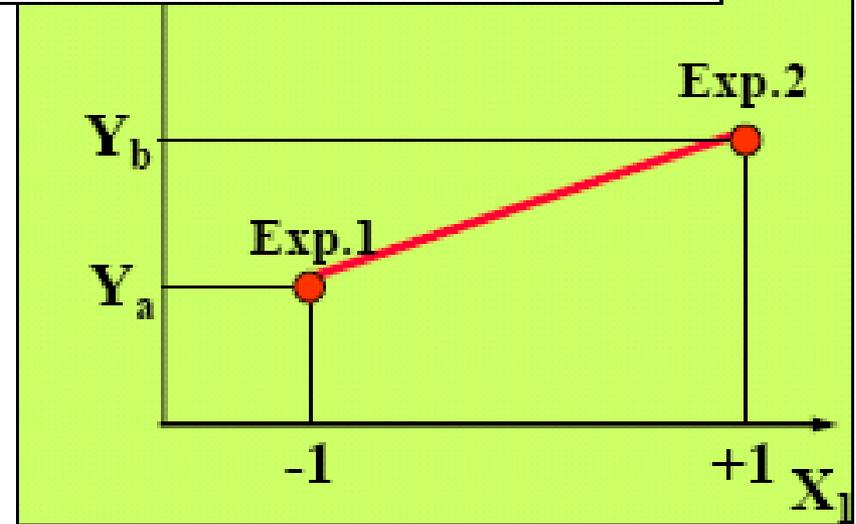
$$Y (\text{réponse}) = b_0 + b_1 X_1 + b_2 X_2 + b_{12} X_1 X_2$$

Factorielles

Poids des Facteurs

**Poids
d'un facteur
(pente)**

| Exp.N° | X_1 |
|--------|-------|
| 1 | -1 |
| 2 | 1 |



Modèle : $Y = b_0 + b_1 X_1$

$$b_1 = \frac{Y_b - Y_a}{2}$$

« poids du facteur X_1 » $b_1 = \frac{E_1}{2}$

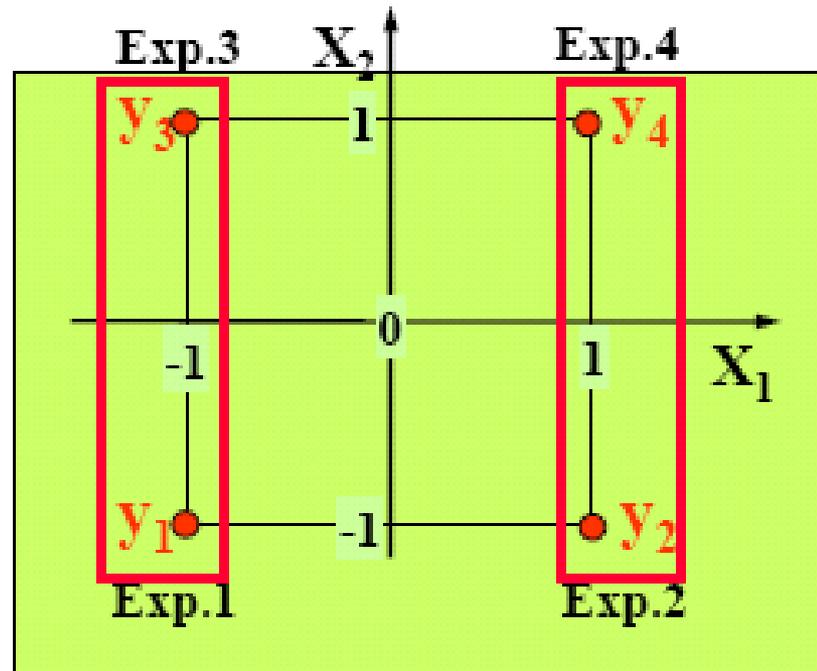
$$b_0 = \frac{Y_b + Y_a}{2}$$

« valeur moyenne » "réponse théorique"
au centre du domaine

Effets des Facteurs

Poids pour deux facteurs

| Exp.N° | X_1 | X_2 |
|--------|-------|-------|
| 1 | -1 | -1 |
| 2 | 1 | -1 |
| 3 | -1 | 1 |
| 4 | 1 | 1 |



$E_1/2 =$ Poids du facteur X_1 :
 $[(y_2+y_4)/2] - [(y_1+y_3)/2]$

Réponse moyenne pour $X_1 = -1$
 $(y_1+y_3)/2$

Réponse moyenne pour $X_1 = 1$
 $(y_2+y_4)/2$

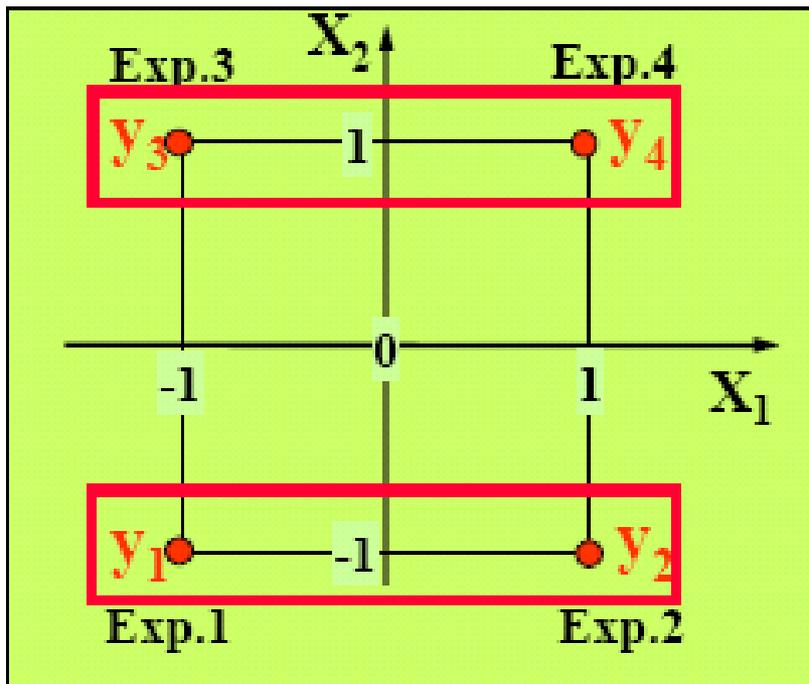


poids de X_1 : $E_1/2 = \frac{-y_1 + y_2 - y_3 + y_4}{2}$

Effets des Facteurs

Poids pour deux facteurs

| Exp.N° | X ₁ | X ₂ |
|--------|----------------|----------------|
| 1 | -1 | -1 |
| 2 | 1 | -1 |
| 3 | -1 | 1 |
| 4 | 1 | 1 |



$E_2 / 2 =$ Poids du facteur X₂:

$$[(y_3 + y_4) / 2] - [(y_1 + y_2) / 2]$$

Réponse moyenne pour X₂ = -1

$$(y_1 + y_2) / 2$$

Réponse moyenne pour X₂ = 1

$$(y_3 + y_4) / 2$$

poids de X₂ :

$$E_2 / 2 = \frac{-y_1 - y_2 + y_3 + y_4}{2}$$

Notion d'interaction

On peut envisager **une autre approche** pour calculer le poids b_i d'un facteur X_i , en tenant compte du niveau d'un **second facteur X_j** :

 appelons $b_1^{(+)}$ le poids du facteur X_1 quand le facteur X_2 est au niveau +1

$$b_1^{(+)} = (y_4 - y_3)/2$$

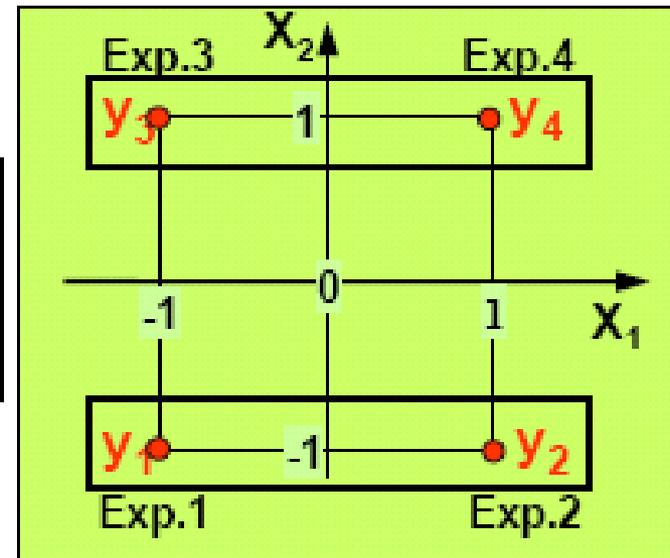
 appelons $b_1^{(-)}$ le poids du facteur X_1 quand le facteur X_2 est au niveau -1

$$b_1^{(-)} = (y_2 - y_1)/2$$

Effets des Facteurs

Poids pour deux facteurs

| Exp.N° | X ₁ | X ₂ |
|--------|----------------|----------------|
| 1 | -1 | -1 |
| 2 | 1 | -1 |
| 3 | -1 | 1 |
| 4 | 1 | 1 |



Poids global de X₁ =
moyenne de b₁⁽⁺⁾ et b₁⁽⁻⁾ :
$$\frac{[(y_4 - y_3)/2] + [(y_2 - y_1)/2]}{2}$$

Poids de X₁
pour X₂ = -1
b₁⁽⁻⁾ = (y₂ - y₁)/2

Poids de X₁
pour X₂ = +1
b₁⁽⁺⁾ = (y₄ - y₃)/2

"poids global" de X₁ : b₁ =
$$\frac{-y_1 + y_2 - y_3 + y_4}{4}$$

Notion d 'interaction

L'effet du facteur X_1 **dépend-il du niveau** auquel se trouve X_2 ?

Il suffit de **comparer $b_1^{(-)}$ et $b_1^{(+)}$** pour répondre à cette question.



Symbolisons par b_{12} **la demi-différence** de ces deux poids partiels :

$$b_{12} = 1/2[b_1^{(+)} - b_1^{(-)}]$$

On appellera b_{12} « l' **interaction** » entre les facteurs X_1 et X_2 .

b_{12} correspond à la **variation de b_1** lorsque X_2 passe de -1 à 1 ($\Delta X_2 = 2$).

$$b_1^{(+)} = (y_4 - y_3)/2$$

$$b_1^{(-)} = (y_2 - y_1)/2$$

$$b_{12} = 1/2[b_1^{(+)} - b_1^{(-)}]$$



$$b_{12} = \frac{+y_1 - y_2 - y_3 + y_4}{4}$$

Nous calculons ce nouveau poids de la même manière que les autres, à l'aide d'une nouvelle combinaison linéaire des y_i divisée par 4.



correspondant à une colonne non encore écrite de la matrice X :

On peut vérifier que cette colonne correspond au produit de la colonne X_1 par la colonne X_2 .

| Exp.N° | X_0 | X_1 | X_2 | $X_2 X_2$ | 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 |

Modèle synergique : deux facteurs

Plan factoriel complet pour 2 facteurs avec interaction

| Exp.N° | X_1 | X_2 |
|--------|-------|-------|
| 1 | -1 | -1 |
| 2 | 1 | -1 |
| 3 | -1 | 1 |
| 4 | 1 | 1 |

Matrice d'expériences

Plan factoriel complet 2^2

Toutes les combinaisons de 2 facteurs
pouvant prendre 2 niveaux

$$\text{Modèle : } Y = b_0 + b_1 X_1 + b_2 X_2 + b_{12} X_1 X_2$$

| Exp.N° | X_0 | X_1 | X_2 | $X_2 X_2$ |
|--------|-------|-------|-------|-----------|
| 1 | 1 | -1 | -1 | 1 |
| 2 | 1 | 1 | -1 | 1 |
| 3 | 1 | -1 | 1 | 1 |
| 4 | 1 | 1 | 1 | 1 |

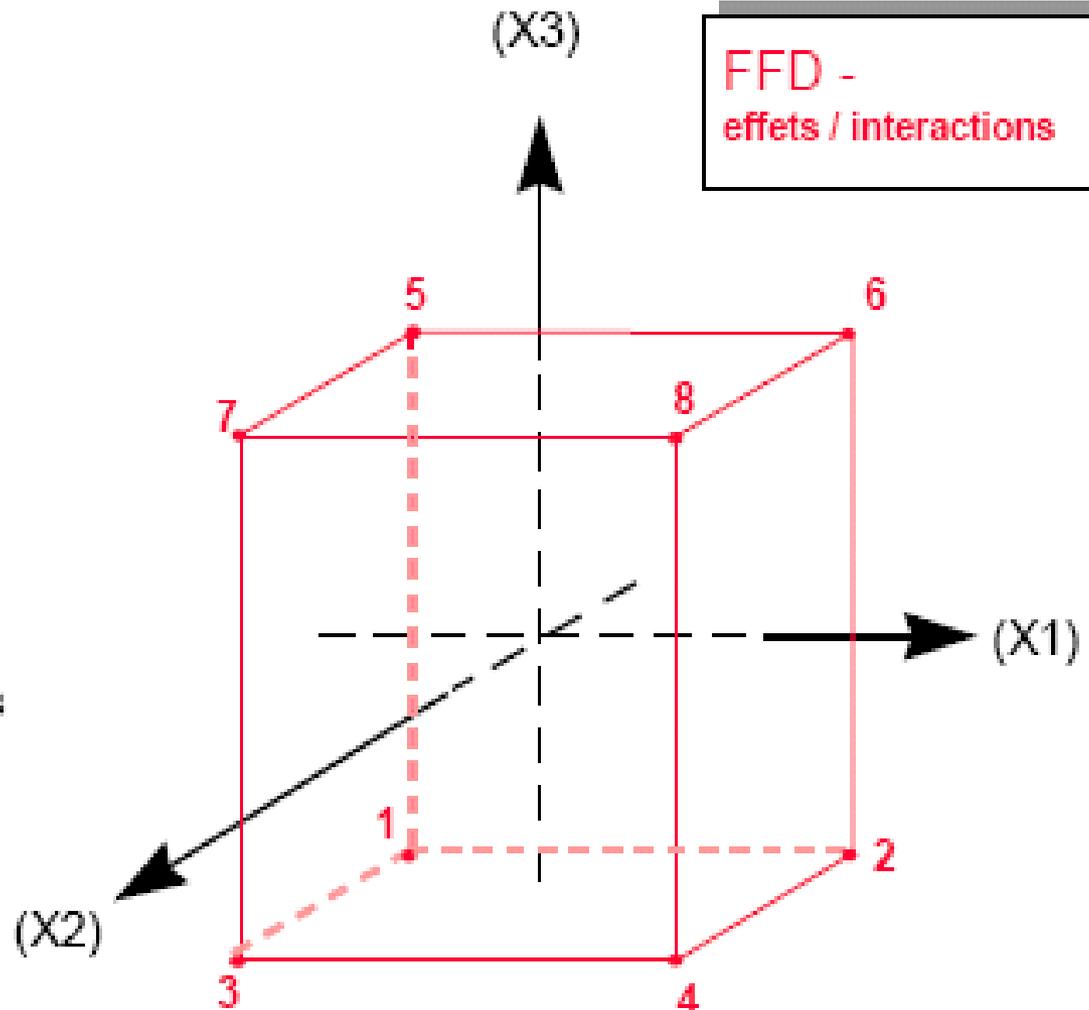
Matrice du modèle

Matrice d'expériences: 3 facteurs

| Essai | X1 | X2 | X3 |
|-------|----|----|----|
| 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 |

Matrice d'expériences
Plan factoriel complet 2^3
Toutes les combinaisons
de 3 facteurs pouvant
prendre 2 niveaux

Algorithme de Yates



Modèle synergique : 3 facteurs

$$Y = b_0 + b_1 X_1 + b_2 X_2 + b_3 X_3 + b_{12} X_1 X_2 + b_{13} X_1 X_3 + b_{23} X_2 X_3 + b_{123} X_1 X_2 X_3$$

1 terme constant

3 effets principaux b_i

1 effet d'interaction du second ordre b_{ijk}

3 effets d'interaction du premier ordre b_{ij}

| N°exp. | X_0 | X_1 | X_2 | X_3 | $X_1 X_2$ | $X_1 X_3$ | $X_2 X_3$ | $X_1 X_2 X_3$ |
|--------|-------|-------|-------|-------|-----------|-----------|-----------|---------------|
| 1 | 1 | -1 | -1 | -1 | 1 | 1 | 1 | -1 |
| 2 | 1 | 1 | -1 | -1 | -1 | -1 | 1 | 1 |
| 3 | 1 | -1 | 1 | -1 | -1 | 1 | -1 | 1 |
| 4 | 1 | 1 | 1 | -1 | 1 | -1 | -1 | -1 |
| 5 | 1 | -1 | -1 | 1 | 1 | -1 | -1 | 1 |
| 6 | 1 | 1 | -1 | 1 | -1 | 1 | -1 | -1 |
| 7 | 1 | -1 | 1 | 1 | -1 | -1 | 1 | -1 |
| 8 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |

Matrice du modèle

Exemple d'utilisation d'une matrice factorielle complète : étude quantitative des facteurs

La réaction enzymatique est réalisée par la culture de la souche microbienne sélectionnée, dans un milieu nutritif contenant le substrat à déshydrogéner :
c'est le développement de la souche qui libère le système enzymatique dans le milieu nutritif et qui "métabolise" le substrat.

| Facteurs | Bornes | |
|--|--------|-----|
| | (-1) | (1) |
| U_1 : Quantité de liqueur de maïs en g/l | 10 | 20 |
| U_2 : Durée de réaction en heures | 24 | 48 |
| U_3 : Quantité de glucose en g/l | 5 | 10 |

RÉPONSE Y étudiée :

Quantité de substrat (mg) **déshydrogénée** enzymatiquement.

Etude d'une déshydrogénation enzymatique

Matrice d'expériences

| Exp.N° | X ₁ | X ₂ | X ₃ | Y |
|--------|----------------|----------------|----------------|-----|
| 1 | -1 | -1 | -1 | 230 |
| 2 | 1 | -1 | -1 | 205 |
| 3 | -1 | 1 | -1 | 110 |
| 4 | 1 | 1 | -1 | 70 |
| 5 | -1 | -1 | 1 | 270 |
| 6 | 1 | -1 | 1 | 220 |
| 7 | -1 | 1 | 1 | 110 |
| 8 | 1 | 1 | 1 | 70 |

Plan d'expérimentation

| U ₁ Quantité liq.maïs | U ₂ Durée réaction | U ₃ Qté glucose | mg subst transf. |
|--|-------------------------------------|----------------------------------|------------------------|
| 10 | 24 | 5 | 230 |
| 20 | 24 | 5 | 205 |
| 10 | 48 | 5 | 110 |
| 20 | 48 | 5 | 70 |
| 10 | 24 | 10 | 270 |
| 20 | 24 | 10 | 220 |
| 10 | 48 | 10 | 110 |
| 20 | 48 | 10 | 70 |

Matrice du modèle

| Exp.N° | X_0 | X_1 | X_2 | X_3 | X_1X_2 | X_1X_3 | X_2X_3 | $X_1X_2X_3$ | Y |
|--------|-------|-------|-------|-------|----------|----------|----------|-------------|-----|
| 1 | 1 | -1 | -1 | -1 | 1 | 1 | 1 | -1 | 230 |
| 2 | 1 | 1 | -1 | -1 | -1 | -1 | 1 | 1 | 205 |
| 3 | 1 | -1 | 1 | -1 | -1 | 1 | -1 | 1 | 110 |
| 4 | 1 | 1 | 1 | -1 | 1 | -1 | -1 | -1 | 70 |
| 5 | 1 | -1 | -1 | 1 | 1 | -1 | -1 | 1 | 270 |
| 6 | 1 | 1 | -1 | 1 | -1 | 1 | -1 | -1 | 220 |
| 7 | 1 | -1 | 1 | 1 | -1 | -1 | 1 | -1 | 110 |
| 8 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 70 |

| | | | | | | | |
|-------|-------|-------|-------|-------|-------|-------|-------|
| I_1 | I_2 | I_3 | I_4 | I_5 | I_6 | I_7 | I_8 |
|-------|-------|-------|-------|-------|-------|-------|-------|

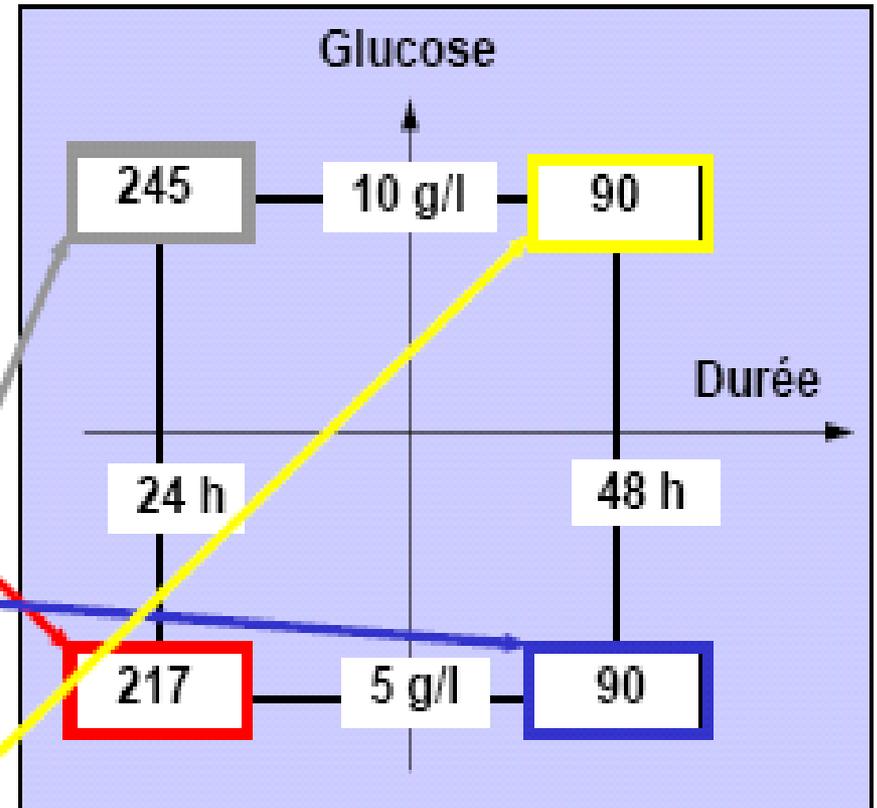
Exploitation des résultats :

- Calcul des coefficients
- Analyse Statistique des résultats
- Significativité des coefficients

Représentation graphique d'une interaction

Matrice d'expériences

| X_1 | X_2 | X_3 | Y |
|-------|-------|-------|-----|
| -1 | -1 | -1 | 230 |
| 1 | -1 | -1 | 205 |
| -1 | 1 | -1 | 110 |
| 1 | 1 | -1 | 70 |
| -1 | -1 | 1 | 270 |
| 1 | -1 | 1 | 220 |
| -1 | 1 | 1 | 110 |
| 1 | 1 | 1 | 70 |



a/ Calcul de coefficients

| Nom | Coefficients |
|-----------|--------------|
| b_0 | 160,6 |
| b_1 | -19,4 |
| b_2 | -70,6 |
| b_3 | 6,9 |
| b_{12} | -0,6 |
| b_{13} | -3,1 |
| b_{23} | -6,9 |
| b_{123} | 3,1 |

La durée de la réaction (facteur X_2) a le plus d'influence : en moyenne le rendement diminue de 140 g/l quand on passe de 24 à 48 h.

Cela signifie que le produit de déshydrogénation est à son tour métabolisé, il ne faut pas laisser la culture trop se développer.

La quantité de liqueur de maïs (facteur X_1) doit également être choisi à son niveau minimum (10 g/l),

La quantité de glucose (facteur X_3) n'a proportionnellement que peu d'effet.

b/ Analyse statistique des résultats

Analyse des coefficients

Le coefficient b_i est distribué selon une **distribution de Student** de **moyenne b_i** , d'**écart-type $e.t.(b_i)$** et $(n-p)$ degrés de liberté.

Si X = matrice d'**Hadamard** :



$$e.t.(b_i) = \frac{s}{\sqrt{n}}$$

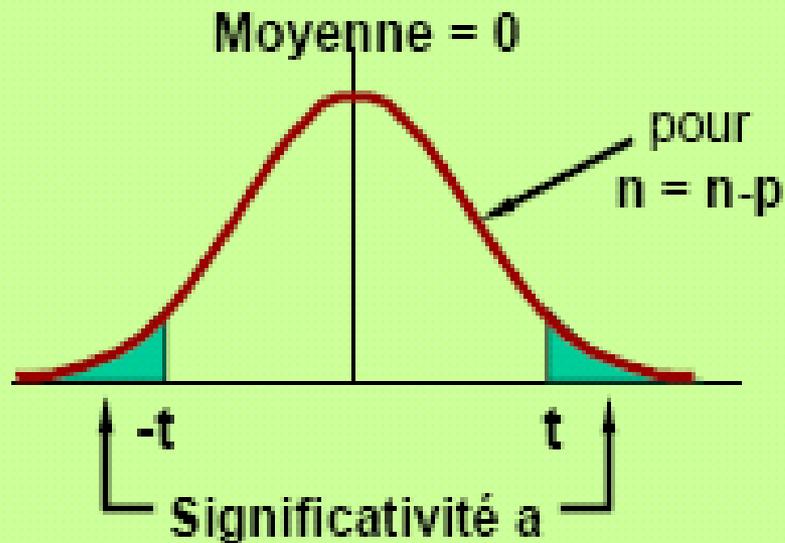
Intervalle de confiance pour b_i :



$$b_i \pm t \ e.t.(b_i)$$

c/ Significativité des coefficients

Loi de Student



$$t = \frac{b_i - b_i^0}{\text{é.t.}(b_i)}$$

$$t = \frac{b_i}{\text{é.t.}(b_i)}$$

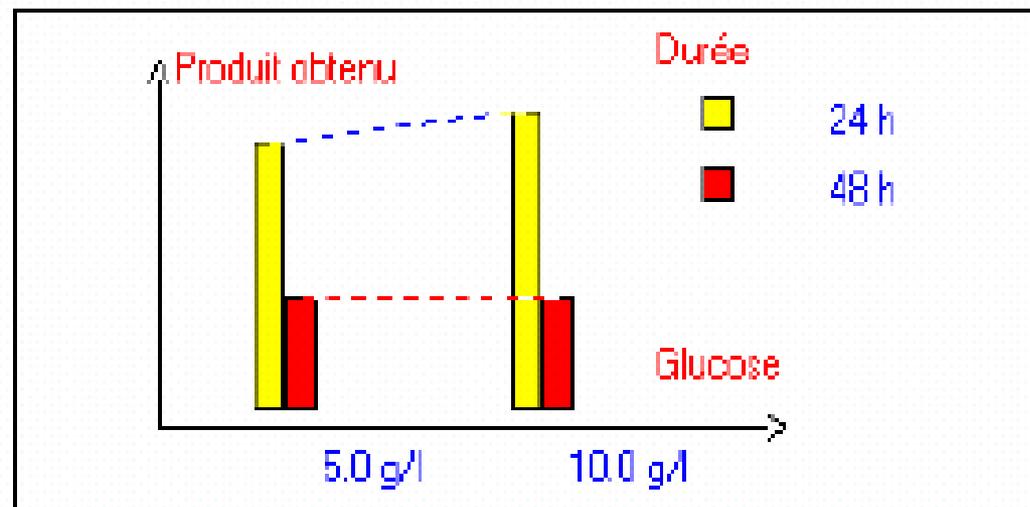
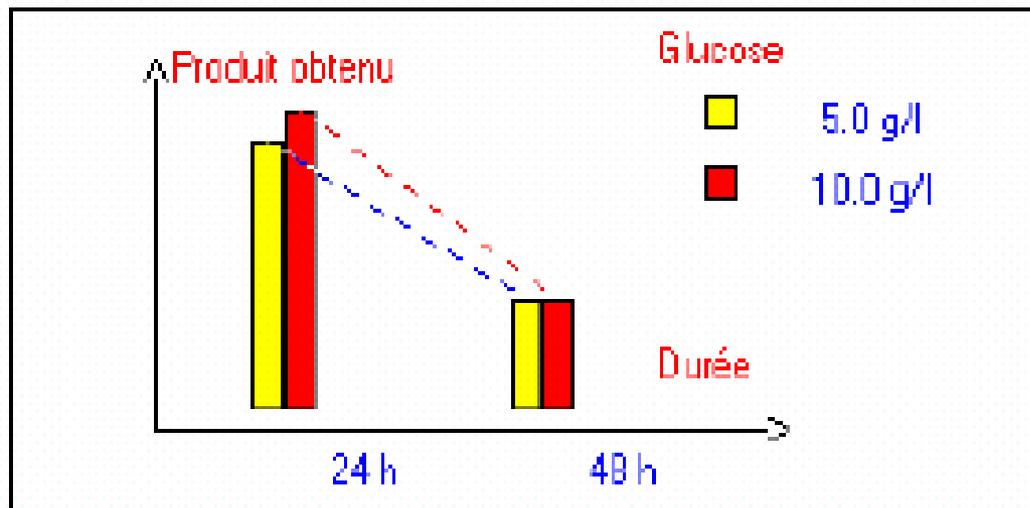
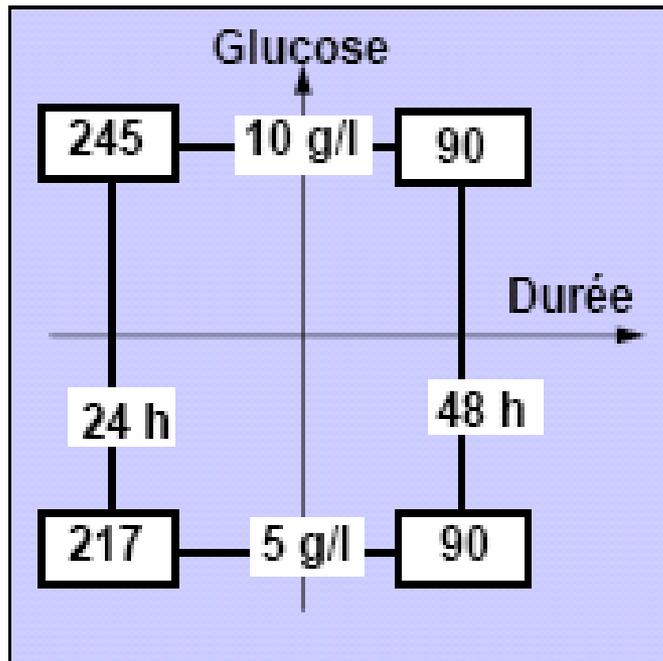
Exploitation des résultats

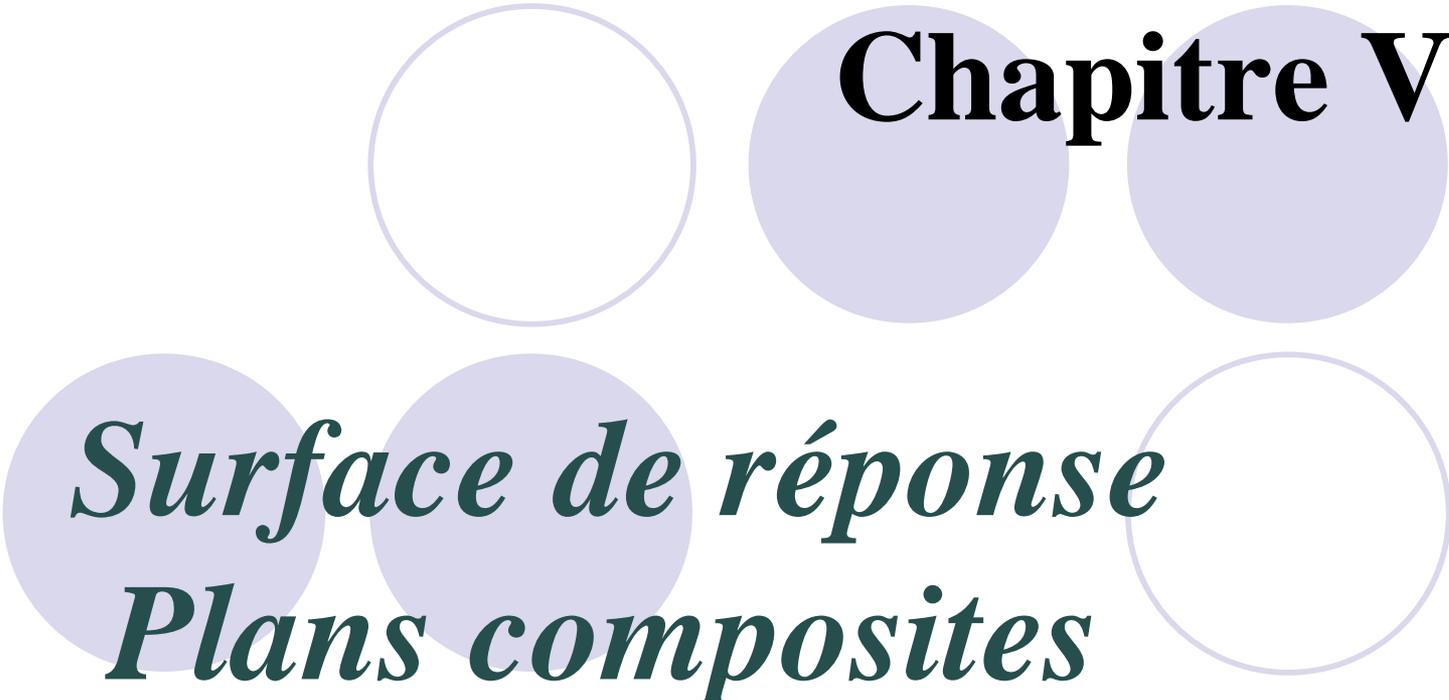
Réexamen «statistique» en négligeant b_{123} comme coefficient :

➔ sa valeur est alors considérée comme **une estimation** de $\frac{\sigma}{\sqrt{N}}$

| Nom | Coefficient | Ecart-type | t.exp. | Signif. % |
|----------|-------------|------------|--------|-----------|
| b_0 | 160,6 | 3.1 | 51.40 | * |
| b_1 | -19,4 | 3.1 | -6.20 | 10.6 |
| b_2 | -70,6 | 3.1 | -22.60 | * |
| b_3 | 6,9 | 3.1 | 2.20 | 27.9 |
| b_{12} | -0,6 | 3.1 | -0.20 | 86.8 |
| b_{13} | -3,1 | 3.1 | -1.00 | 50.0 |
| b_{23} | -6,9 | 3.1 | -2.20 | 27.9 |

Représentation graphique d'une interaction





Chapitre V

Surface de réponse

Plans composites

PCCD

MATRICES D'EXPERIENCES et METHODOLOGIE

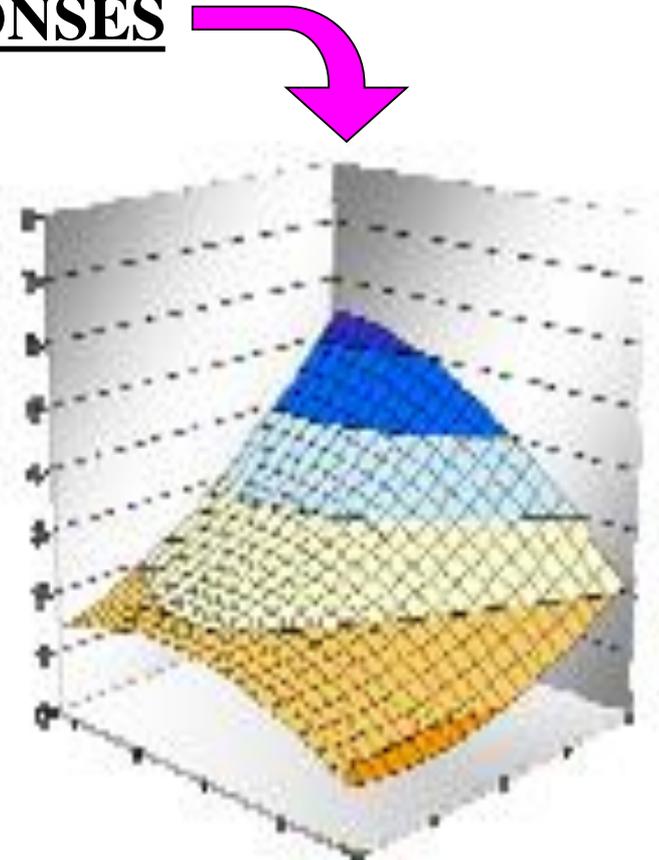
Objectifs :

- Études quantitatives des réponses : **modélisation** prévisionnelle du phénomène étudié,
- **Optimisation** : SURFACES de REPONSES

Le modèle quadratique par rapport aux variables explicatives x_i ; établi à l'aide d'un PCC :
$$Y = b_0 + \sum_i b_i X_i + \sum_{i,j} b_{ij} X_i X_j + \sum_i b_{ii} X_i^2$$

On aura à estimer des termes appartenant à 4 familles :

- b_0 : terme constant
- b_i : terme de premier degré
- b_{ii} : terme carré (du deuxième degré)
- b_{ij} : terme rectangle (entre X_i et X_j)



Dans la méthodologie des surfaces de réponses; les plans composites permettent d'obtenir la meilleure précision possible sur la modélisation des résultats.

I/ Construction des plans composites centrés

I-1/ Démarche séquentielle

Les plans PCC permettent une modélisation du second degré. La démarche séquentielle est constituée de trois parties : plan factoriel; plan en étoile; points (plan) au centre.

A/ Plan factoriel

Le plan factoriel combine des facteurs à 2 niveaux notés +1 et -1. Il permet :

- *l'analyse des interactions;
- *de déterminer le modèle

Le nombre d'essais du plan factoriel :

$$N_F = 2^k$$

(k : facteurs, 2: niveaux (-1 et +1))

B/ Plan en étoile

le plan en étoile représente deux essais par facteur. Chaque facteur prend au total 5 niveaux différents dans le plan.

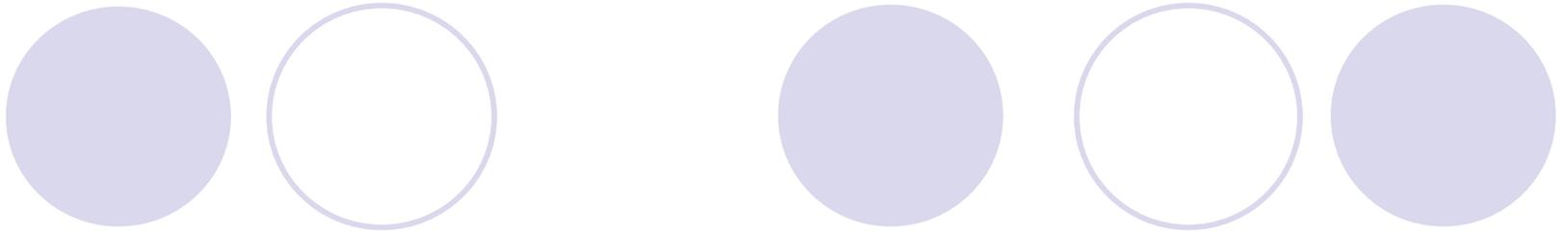
Le nombre d'essais du plan en étoile : $N_a = 2.k$

C/ Points au centre

Ces points peuvent être la solution courante que l'on cherche à améliorer

Ils ont plusieurs rôles :

- Tester la validité du modèle du premier degré;
- S'assurer qu'il n'y a pas de glissement entre deux séries d'essais  stabilité;
- Obtenir une estimation de l'erreur expérimentale;
- Diminuer l'erreur de prédiction près du point central.



Le nombre d'essais des points au centre :

Les points au centre du domaine expérimental (N_0) sont répétés plusieurs fois (?) (Voir ultérieurement).

Le nombre d'essais au centre : N_0

Le nombre total d'essais à réaliser dans le PCC :

$$N = N_F + N_a + N_0$$

I-2/ Exemple de plan PCC avec deux facteurs :

La matrice du plan d'expériences pour $k = 2$:

⇒ 5 niveaux : $-\alpha, -1, 0, +1, +\alpha$ (α à déterminer).

⇒ 3 niveaux : $-1, 0, +1$ si $\alpha = 1$

Le nombre total d'essais à réaliser pour le PCC :

$$N = 2^k + 2.k + N_0$$

$$K = 2 \quad ; \quad \Rightarrow N = 2.^2 + 2.2 + N_0 = 8 + N_0$$

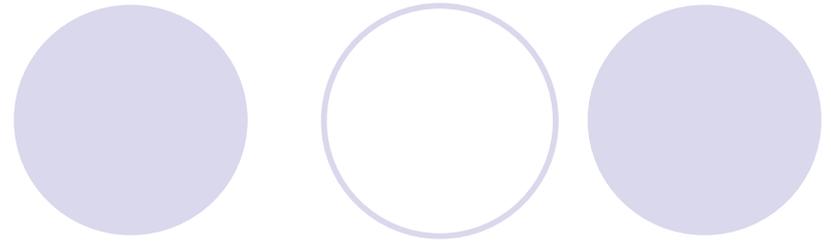
Si $N_0 = 1$; le modèle quadratique au degré 2 aura

$(k+1)(k+2)/2$ variables explicatives (facteurs).

Matrice du plan d'expériences s'écrit :

Matrice d'expériences

| Essais | x_1 | x_2 |
|--------|-----------|-----------|
| 1 | -1 | -1 |
| 2 | +1 | -1 |
| 3 | -1 | +1 |
| 4 | +1 | +1 |
| 5 | $-\alpha$ | 0 |
| 6 | $+\alpha$ | 0 |
| 7 | 0 | $-\alpha$ |
| 8 | 0 | $+\alpha$ |
| 9 | 0 | 0 |
| . | . | . |
| N | 0 | 0 |



Plan factoriel
($N_F = 2^2 = 4$)



Plan en étoile (axial)
($N_a = 2.2 = 4$)



Points au centre (N_0)

● Pour K variables $\implies N_F = 2^K$: termes linéaires en X_i et aux interactions $X_i X_j$.

➤ Les coordonnées des points ont toutes la valeur 1 en valeur absolue. Et chaque point représente un essai à exécuter .

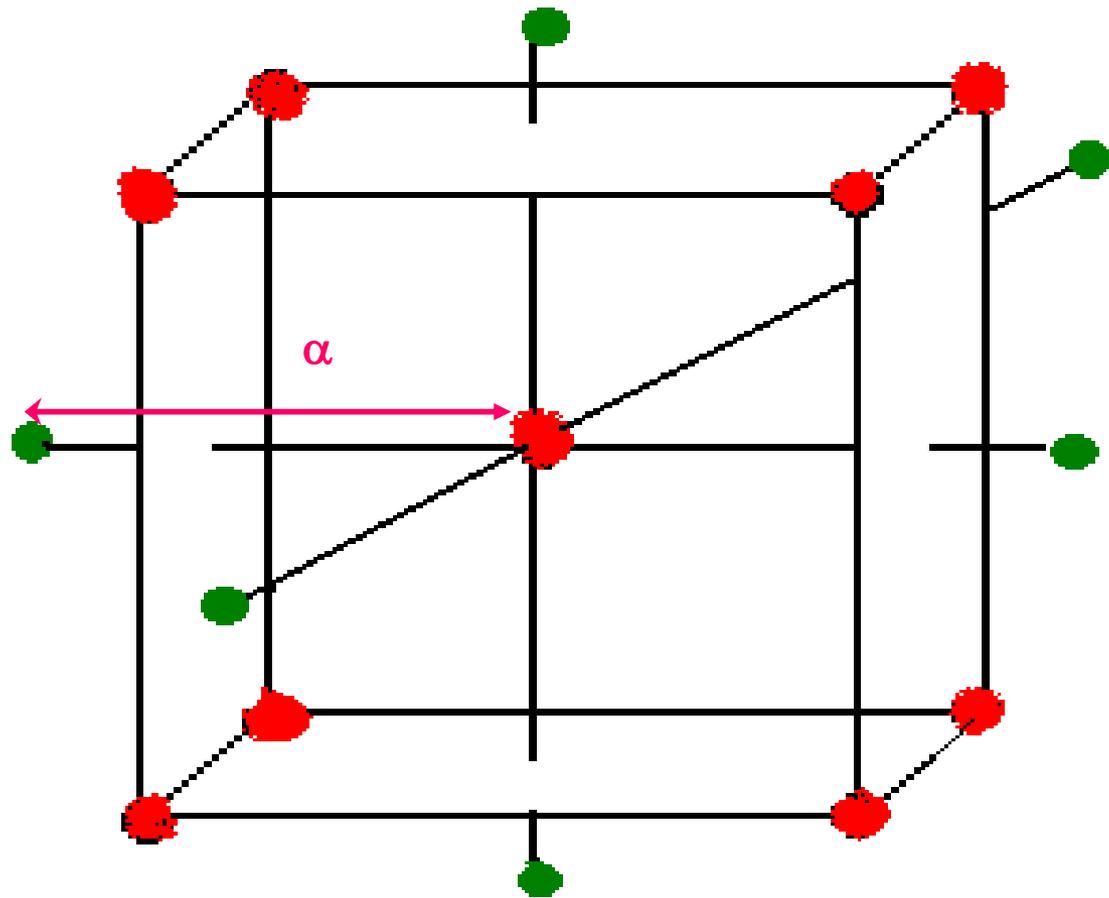
➤ Les points représentatifs sont situés aux sommets d'un cube centré sur l'origine $(0, 0, \dots, 0)$ et dont les faces sont parallèles aux plans de coordonnées à la distance ± 1 (points en rouge)

➤ $N_a = 2.K$ les points expérimentaux sont sur les axes situés à la même distance du centre du domaine (points axiaux en vert).

➤ Les points axiaux sont situés à la même distance $\pm \alpha$ (α calculé à partir du critère de l'orthogonalité).

➤ Les points au centre (N_0) :

- Précision uniforme
- Isovariance par rotation



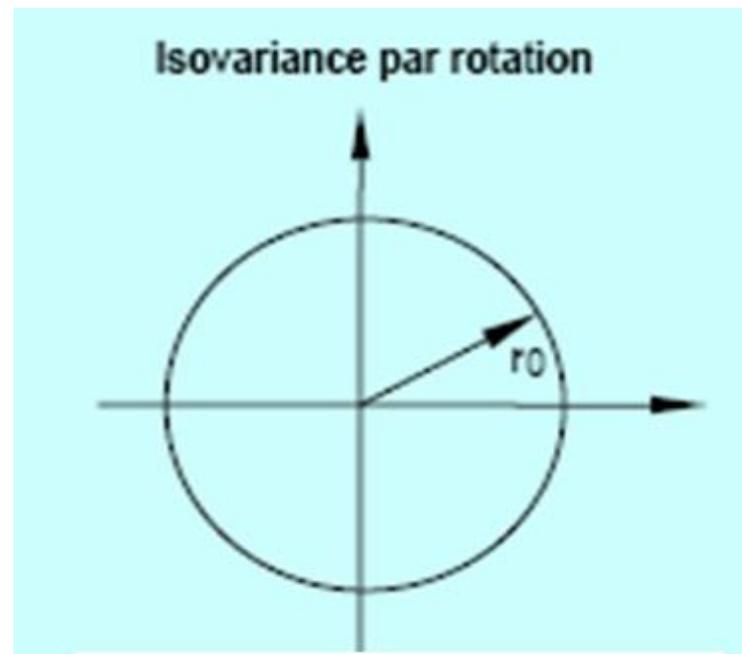
P.C.C. à 3 variables

Critères d'optimalité :

Dans le cas où tous les points axiaux sont situés à la même distance du centre du domaine d'étude, la valeur de α ne sera pas la même (suivant les critères d'optimalité).

a/ Isovariance par rotation

La valeur de α est choisie pour que la variance de la réponse prévue par le modèle ne dépend que de la distance du centre du domaine où tous les points situés sur un même cercle ont même variance.



Dans ce cas, les éléments de la matrice $(X'X)$ doivent satisfaire la relation : $\alpha = (N_F)^{1/4}$

Donc il faut placer les points en étoiles à une distance :

$$\alpha = (N_F)^{1/4}$$

Exemple : Pour le cas d'un PCC à 2 facteurs :

$\alpha = 1,414$ du centre.

● La valeur de α est donc fonction du nombre de points au centre N_0

b/ Précision uniforme

La valeur de α est choisie pour que la variance de la réponse prévue par le modèle ne dépend que de la distance du centre du domaine où tous les points situés sur un même cercle et à l'intérieur ont même variance.

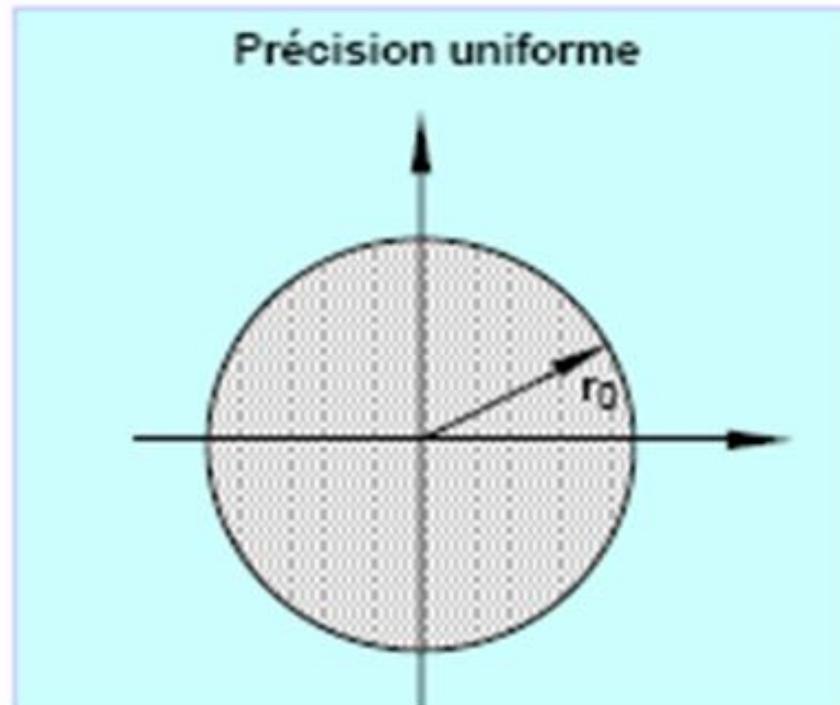


Tableau : Valeur de α en fonction du nombre de points centraux

| | | | | | | | |
|--|--------|--------|---------|-----------|----------|-----------|----------|
| Nb k de facteurs | 2 | 3 | 4 | 5 | 5 | 6 | 6 |
| Matrice factorielle | 2^2 | 2^3 | 2^4 | 2^{5^1} | 2^5 | 2^{6^1} | 2^6 |
| Isovariance par rotation $\alpha = (N_F)^{1/4}$ | 1,414 | 1,682 | 2 | 2 | 2,378 | 2,378 | 2,826 |
| Points factoriels N_F | 4 | 8 | 16 | 16 | 32 | 32 | 64 |
| Points axiaux N_a | 4 | 6 | 8 | 10 | 10 | 12 | 12 |
| Points au centre N_o pour : ▪Orthogonalité ▪Précision uniforme | 8 5 | 9 6 | 12 7 | 10 6 | 17 10 | 15 9 | 24 15 |
| N total d'expériences | 16 13 | 23 20 | 36 31 | 36 32 | 59 52 | 59 53 | 100 91 |

I-3/ Détermination des paramètres du plan

- Le modèle quadratique par rapport aux variables explicatives x_i ; établi à l'aide d'un PCC comportant $\frac{(k+1)(k+2)}{2}$ coefficients :

- $$Y = \underbrace{b_0}_{\text{cste}} + \underbrace{b_1x_1 + b_2x_2}_{\text{linéaire}} + \underbrace{b_{12}x_1x_2}_{\text{interaction}} + \underbrace{b_{11}x_1^2 + b_{22}x_2^2}_{\text{quadratique}}$$
- 

P. 1^{er} degré



P. 2^{ème} degré

Plus généralement :

$$Y = b_0 + \sum_i b_i X_i + \sum_{i,j} b_{ij} X_i X_j + \sum_i b_{ii} X_i^2$$

Modèle du second degré

La réponse Y est estimée à l'aide d'un **polynôme de degré 2**,

par exemple

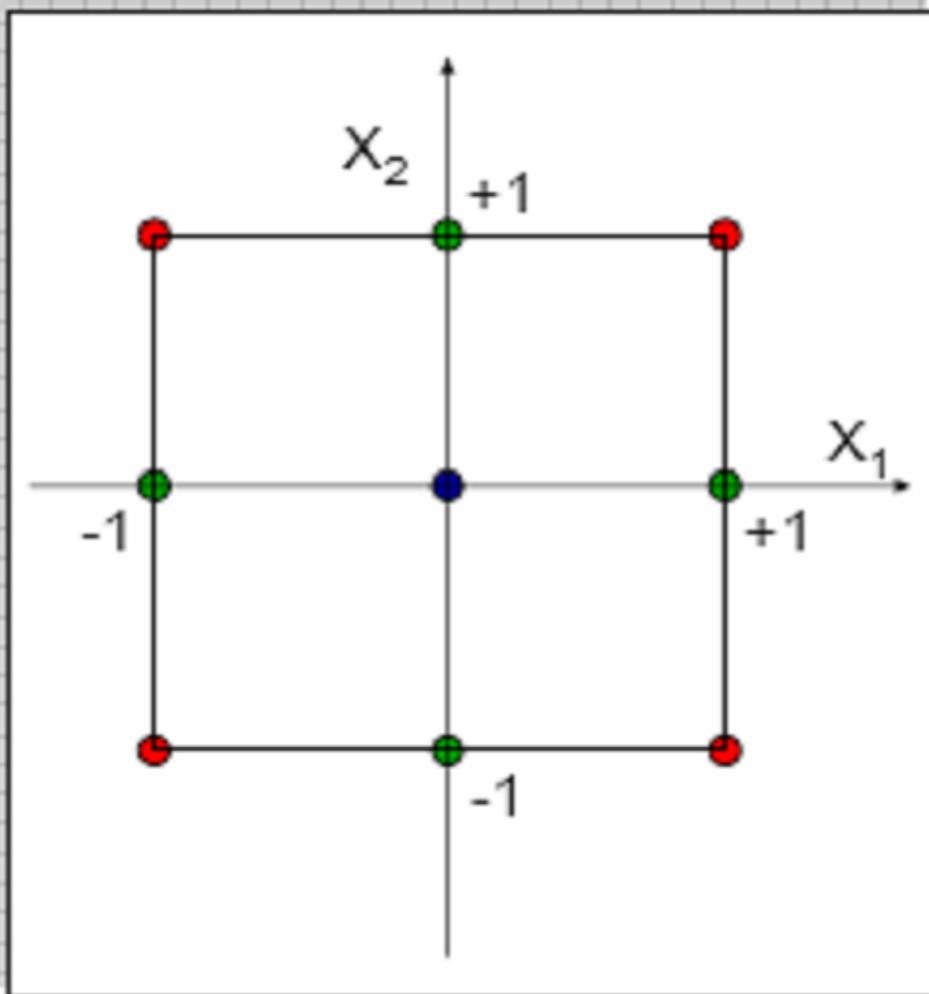
dans le cas de **2 facteurs** il faut estimer **6 coefficients** :

$$Y = b_0 + b_1 X_1 + b_2 X_2 + b_{11} X_1^2 + b_{22} X_2^2 + b_{12} X_1 X_2$$

pour **3 facteurs**, **10 coefficients** :

$$Y = b_0 + b_1 X_1 + b_2 X_2 + b_3 X_3 + b_{11} X_1^2 + b_{22} X_2^2 + b_{33} X_3^2 \\ + b_{12} X_1 X_2 + b_{13} X_1 X_3 + b_{23} X_2 X_3$$

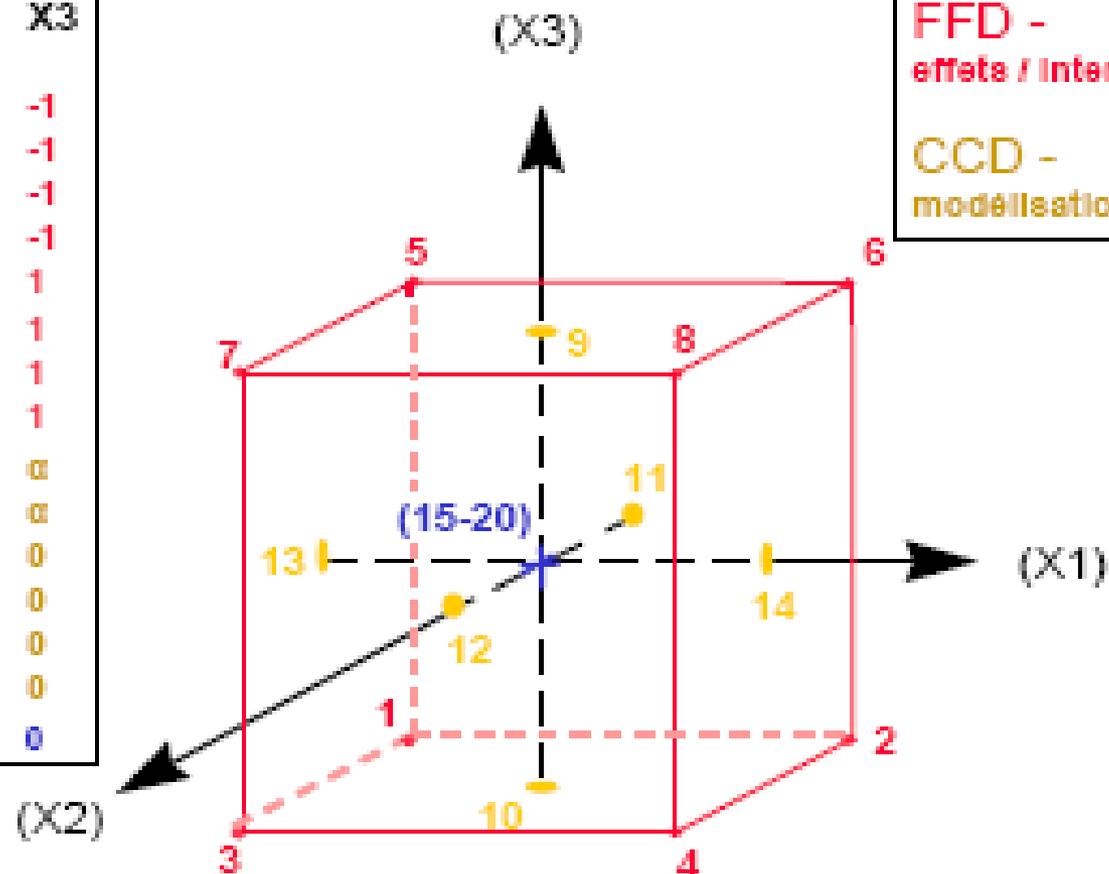
Matrice Composite Face-centrée



| | X1 | X2 |
|---|----|----|
| ● | -1 | -1 |
| ● | +1 | -1 |
| ● | -1 | +1 |
| ● | +1 | +1 |
| ● | -1 | 0 |
| ● | +1 | 0 |
| ● | 0 | -1 |
| ● | 0 | +1 |
| ● | 0 | 0 |
| ● | 0 | 0 |
| ● | 0 | 0 |
| ● | 0 | 0 |
| ● | 0 | 0 |

Plan central composite

| Essai | X1 | X2 | X3 |
|-------|----------|----------|----------|
| 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 |
| 9 | 0 | 0 | α |
| 10 | 0 | 0 | α |
| 11 | 0 | α | 0 |
| 12 | 0 | α | 0 |
| 13 | α | 0 | 0 |
| 14 | α | 0 | 0 |
| 15-20 | 0 | 0 | 0 |



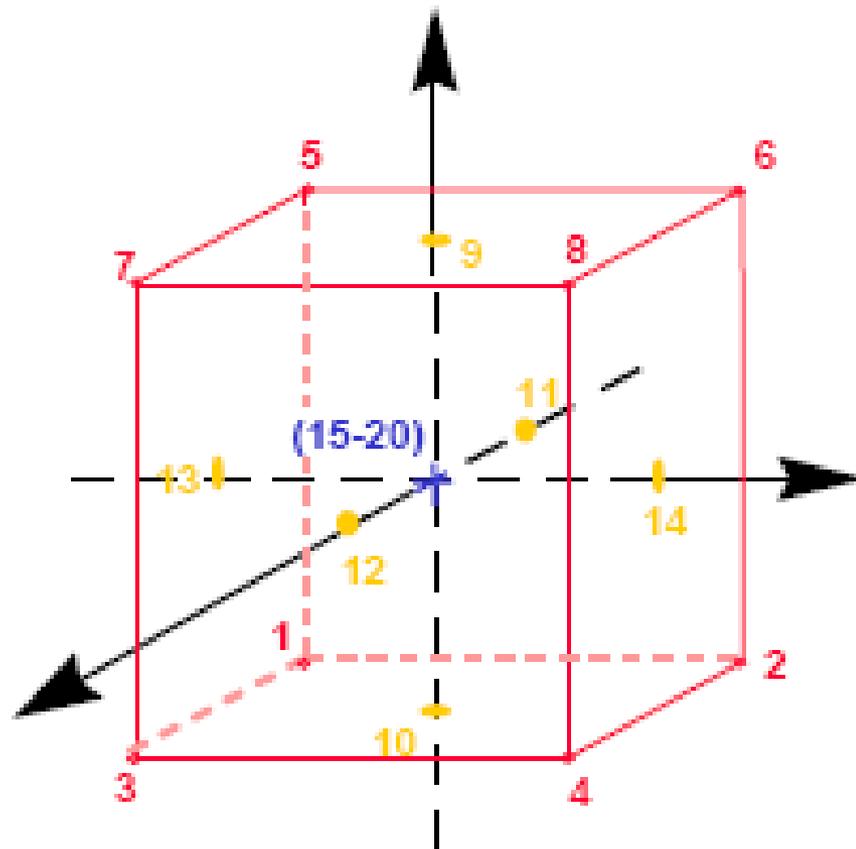
FFD -
effets / interactions

CCD -
modélisation

Composite équiradiale : $\alpha = 1.682$

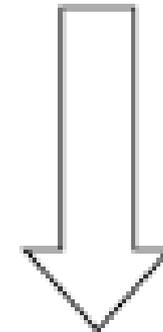
Composite centrée : $\alpha = 1$

Plan central composite



Composite centrée :

$$\alpha = 1$$



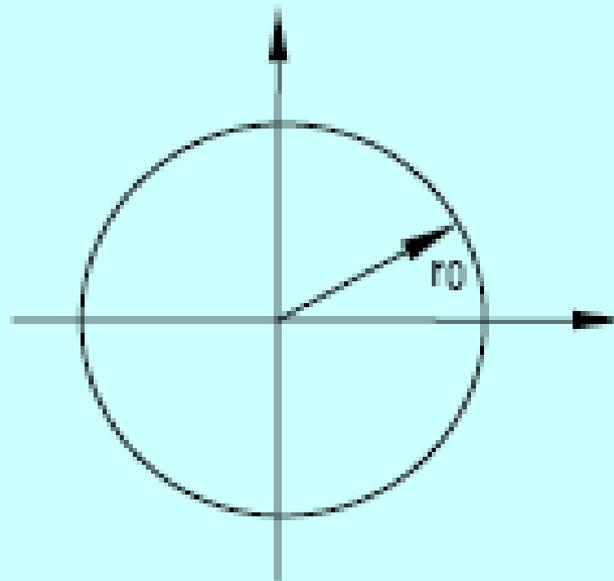
Composite équiradiale :

$$\alpha = 1.682$$

Modèle du second degré

Critères statistiques

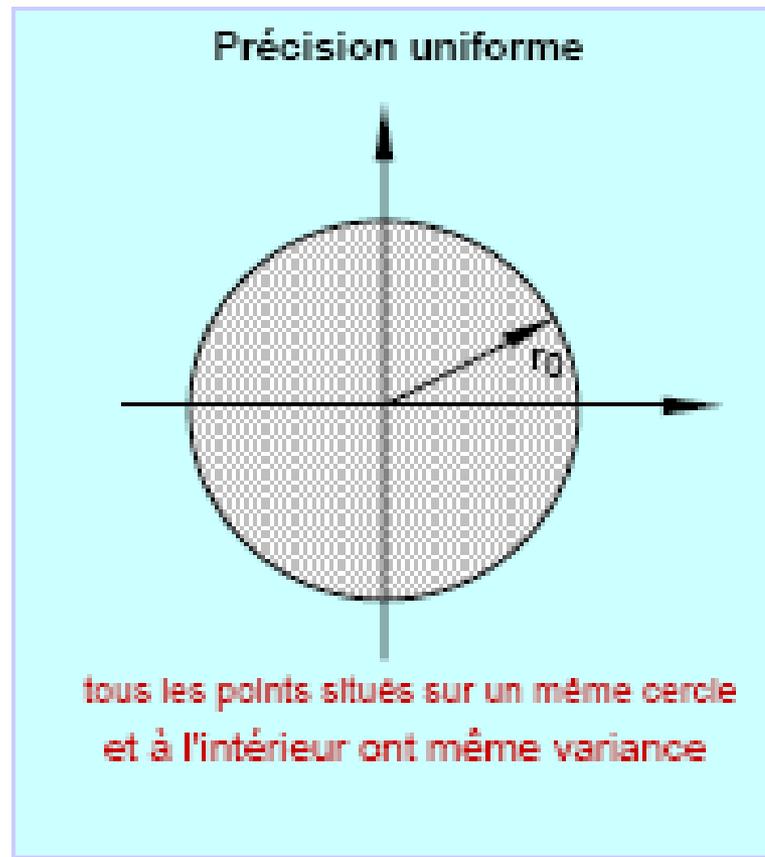
Isovariance par rotation



tous les points situés sur un
même cercle ont même variance

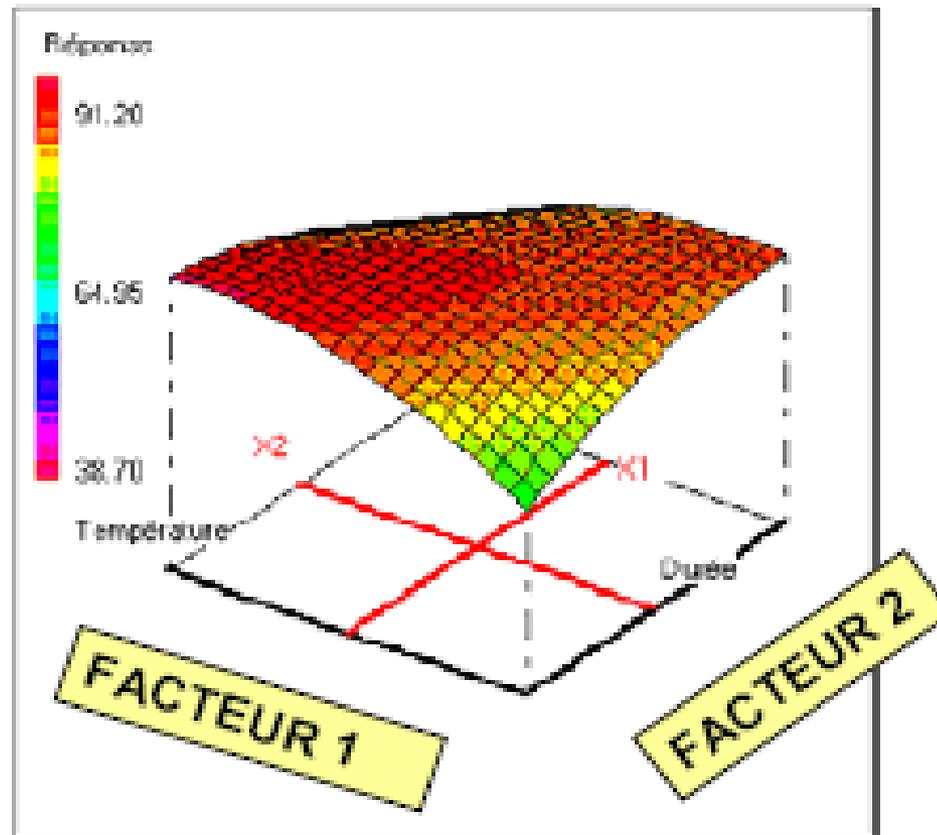
Modèle du second degré

Critères statistiques

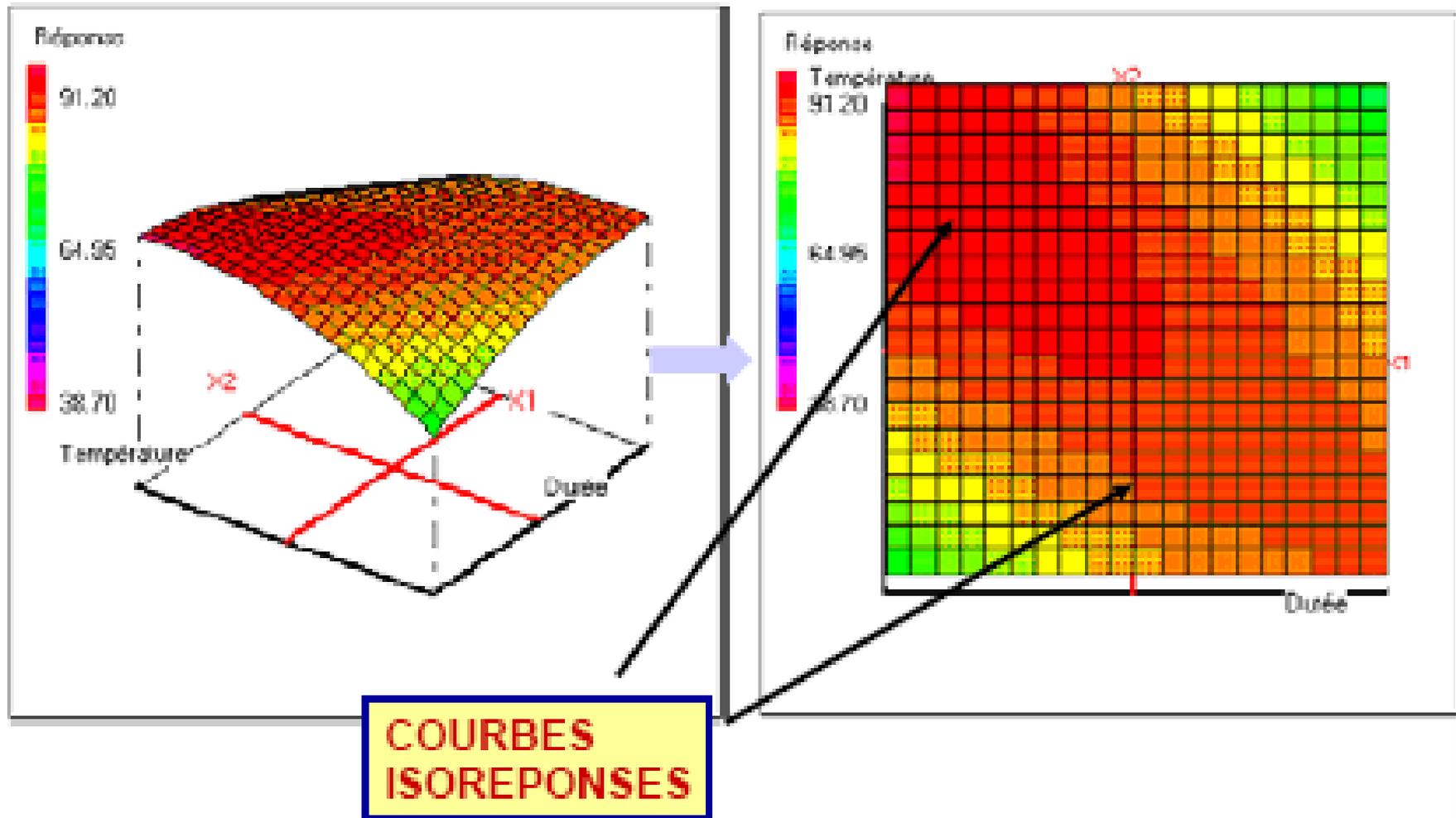


Surface de réponse 3D

R
E
P
O
N
S
E
M
S
O
N
S



Surface de réponse : projection 3D – 2D

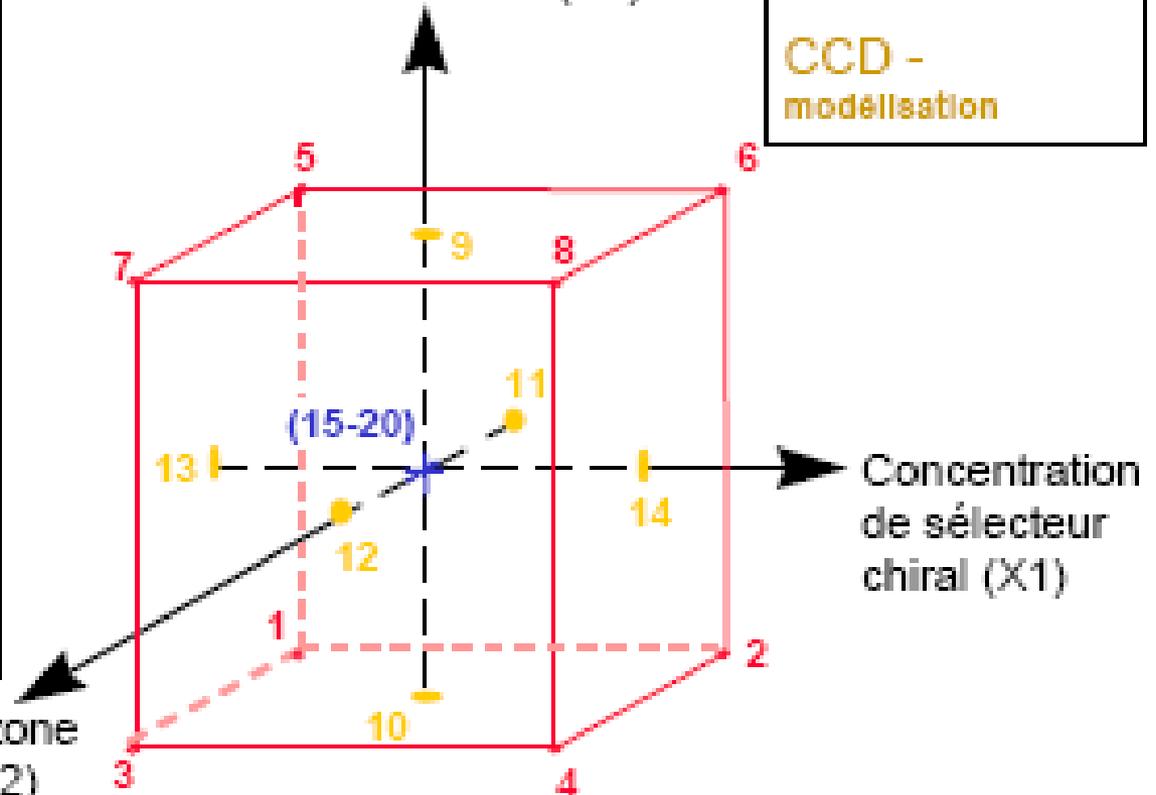


Plan central composite

| Essai | X1 | X2 | X3 |
|-------|----|----|----|
| 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 |
| 9 | 0 | 0 | -1 |
| 10 | 0 | 0 | 1 |
| 11 | 0 | -1 | 0 |
| 12 | 0 | 1 | 0 |
| 13 | -1 | 0 | 0 |
| 14 | 1 | 0 | 0 |
| 15-20 | 0 | 0 | 0 |

Longueur de la zone de séparation (X2)

Pression du gaz de nébulisation (X3)



FFD -
effets / interactions

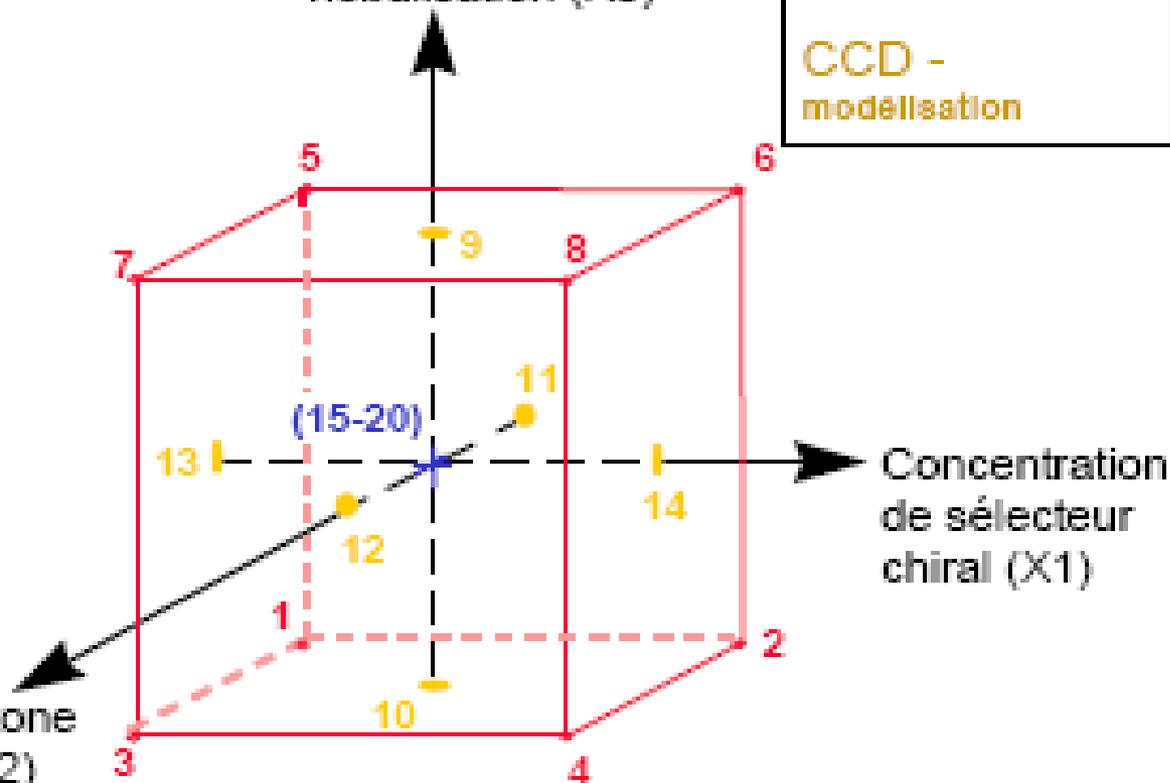
CCD -
modélisation

Plan central composite

| Essai | X1 | X2 | X3 |
|-------|----|----|----|
| 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 |
| 9 | 0 | 0 | -1 |
| 10 | 0 | 0 | 1 |
| 11 | 0 | -1 | 0 |
| 12 | 0 | 1 | 0 |
| 13 | -1 | 0 | 0 |
| 14 | 1 | 0 | 0 |
| 15-20 | 0 | 0 | 0 |

Longueur de la zone de séparation (X2)

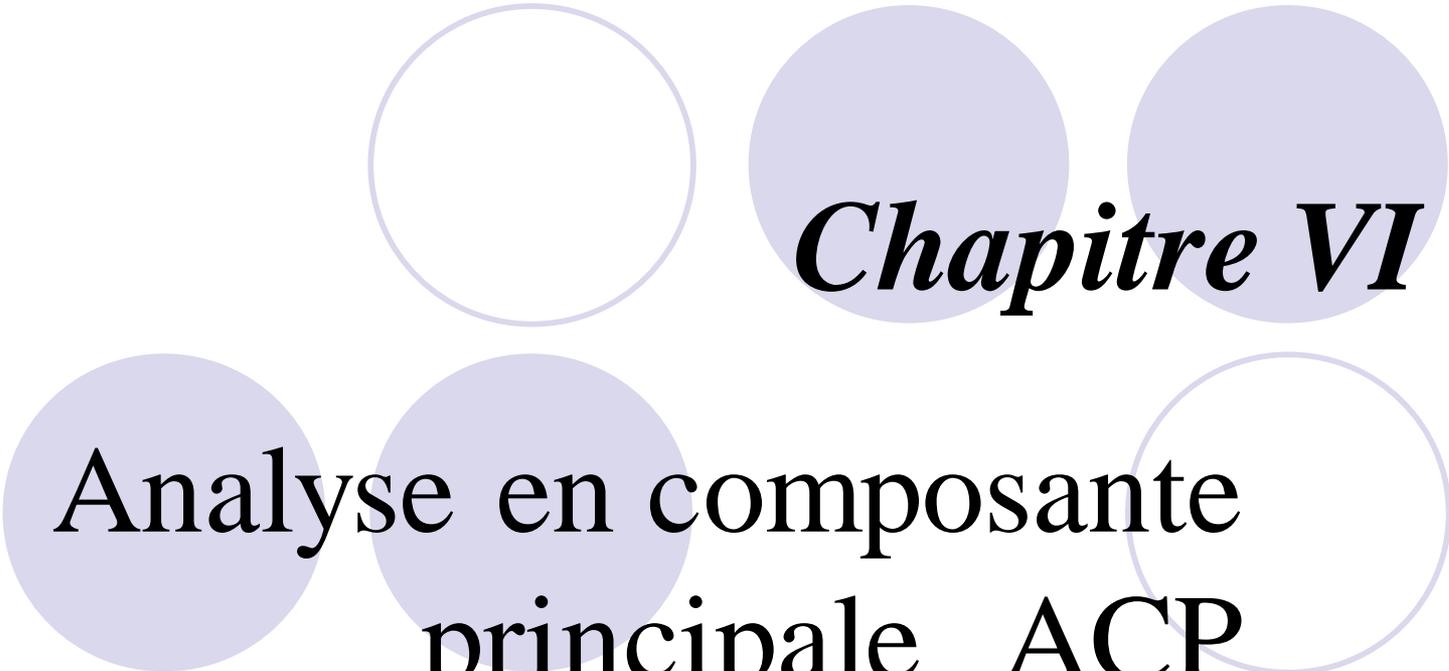
Pression du gaz de nébulisation (X3)



FFD -
effets / interactions

CCD -
modélisation

Concentration de sélecteur chiral (X1)



Chapitre VI

Analyse en composante principale ACP

La méthode de l'ACP

Analyse factorielle en
Composantes **P**incipales

Variables explicatives
et à expliquer **séparées** :

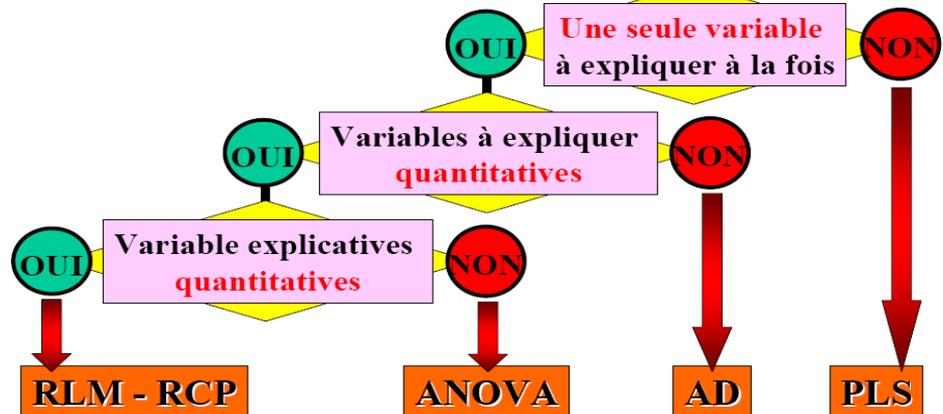
OUI

NON

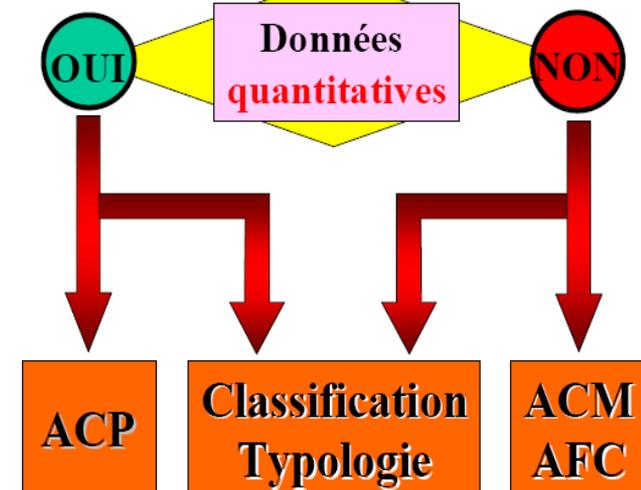
Méthodes **explicatives**

Méthodes **descriptives**

Méthodes **explicatives** :



Méthodes **Descriptives** :



Problème de la Compression des Données

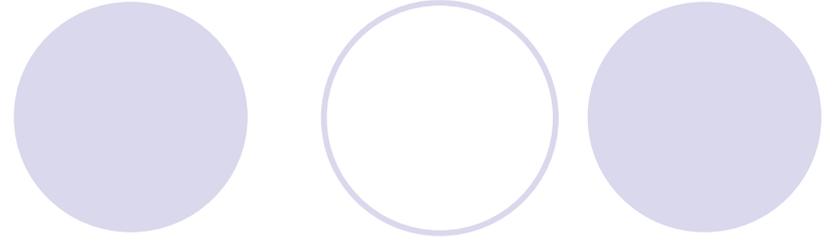
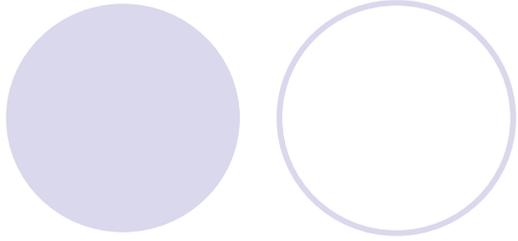
Comment
prendre en compte :

Un **grand nombre** de variables

Un **nombre limité** d'individus

l'information pertinente étant noyée dans une
masse d'informations partiellement **redondantes**
et polluées par **toutes sortes d'erreurs** ?

Peut-on **réduire la masse** des données
sans perdre trop d'information ?
et **améliorer la lisibilité** du contenu des données ?

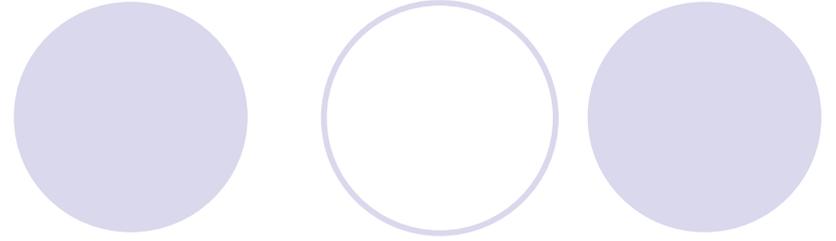
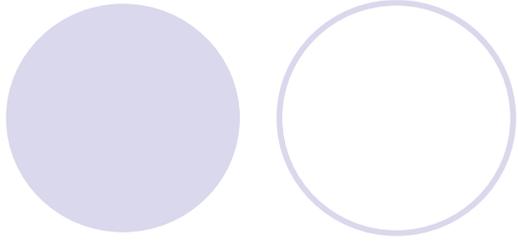


L'Analyse en Composantes Principales (ou Principal Component Analysis)

Elle consiste à **transformer** les **m** variables quantitatives initiales toutes plus ou moins **corrélées** entre elles

en **m** nouvelles variables quantitatives **non corrélées** appelées **composantes principales**.

Puis, si possible, **ne conserver qu'un nombre k réduit** de composantes.



Principe de l'ACP

A partir de la matrice X des données de départ on construit la **matrice M de corrélation** :

$$M = 1/n (X'X)$$

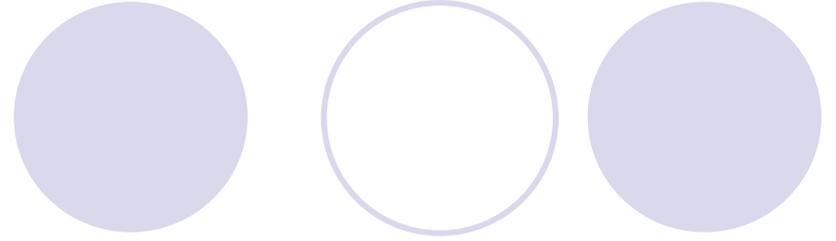
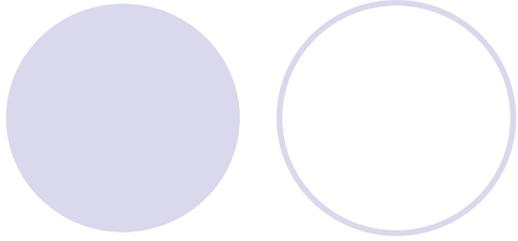
Avec X' matrice transposée de X ;

N.B. : X peut être préalablement transformée en variables centrées, en variables centrées réduites.

La **matrice M de corrélation** est **diagonalisée** pour fournir :

une matrice diagonale D des **valeurs propres** $\lambda_1, \lambda_2, \dots, \lambda_m$

une matrice P des **vecteurs propres**.



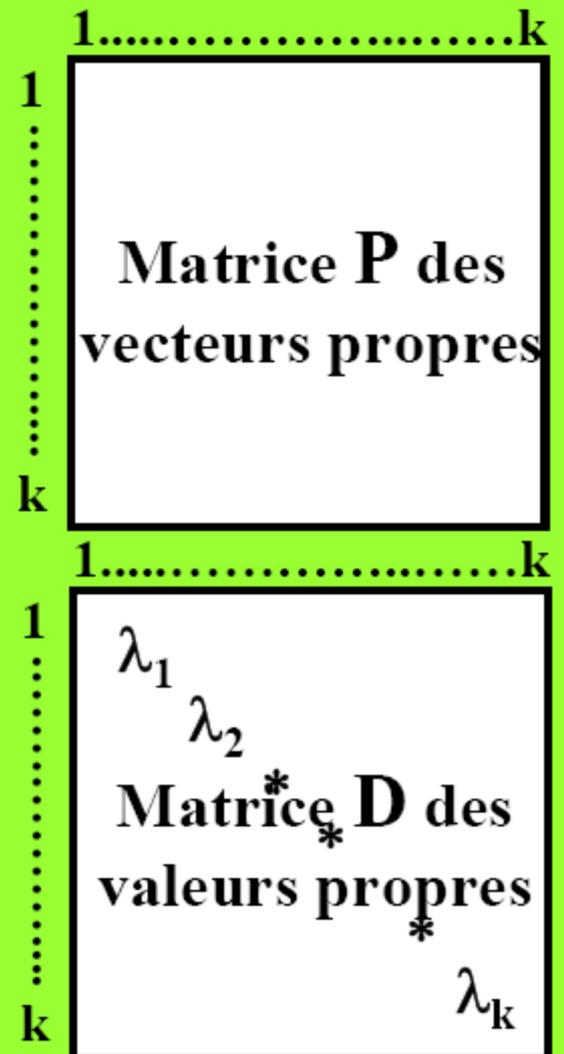
Les **vecteurs colonnes** de la matrice **P** représentent les **combinaisons linéaires** des variables de départ conduisant aux **nouveaux axes factoriels** :



les Composantes Principales.

Les **valeurs propres** $\lambda_1, \lambda_2, \dots, \lambda_m$ représentent les **variances des individus** sur les composantes principales correspondantes.

$$D = P^{-1} M P$$

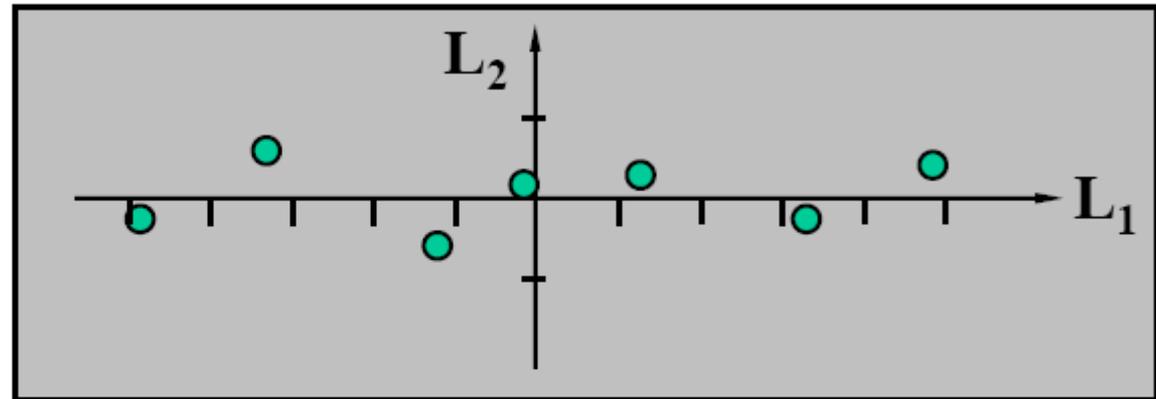


Le rapport de chaque valeur propre à la somme de toutes les valeurs propres donne **la part de toute l'information initiale** visible sur chaque axe :

$$\% \text{ information} = \frac{\lambda_i}{\sum \lambda_i}$$

On peut alors faire une "projection" d'un espace à k dimensions initiales à un espace plus réduit (2, 3) **en sachant quel pourcentage de l'information est conservé** (90 % par exemple).

| L_1 | L_2 |
|--------|--------|
| -2.124 | 0.101 |
| -1.420 | -0.274 |
| -0.569 | 0.249 |
| -0.024 | -0.057 |
| 0.570 | -0.084 |
| 1.417 | 0.174 |
| 2.149 | -0.109 |

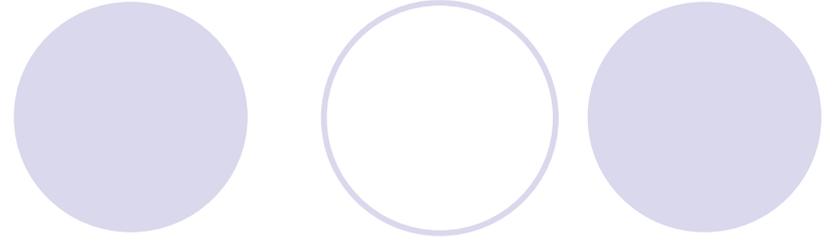
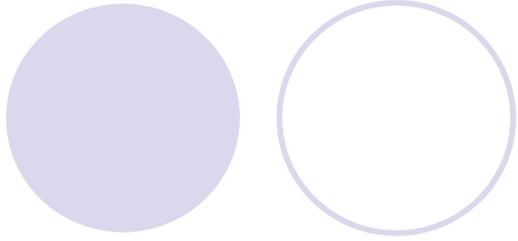


Les descripteurs dans ce nouveau système d'axes sont **totallement décorrélés** :



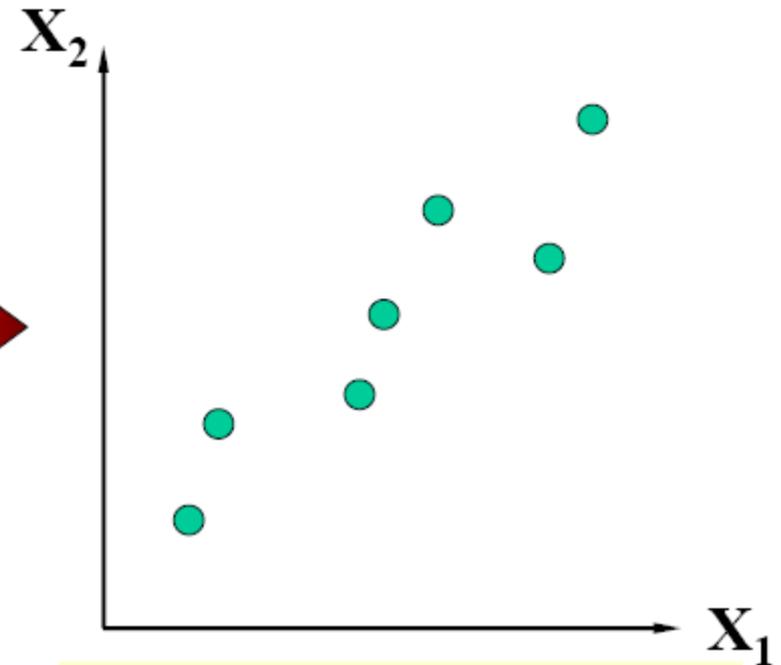
Matrice de
corrélacion

| | L_1 | L_2 |
|-------|-------------|-------|
| L_1 | 1 | |
| L_2 | 0.00 | 1 |



Composantes Principales dans le cas de 2 variables

| X_1 | X_2 |
|-------|-------|
| 2.8 | 13.3 |
| 3.9 | 16.6 |
| 8.5 | 17.6 |
| 9.3 | 20.2 |
| 11.2 | 22.1 |
| 14.9 | 23.9 |
| 16.4 | 27.0 |



Soit 7 individus décrits par deux variables X_1 et X_2 fortement corrélées



Matrice de **corrélation**

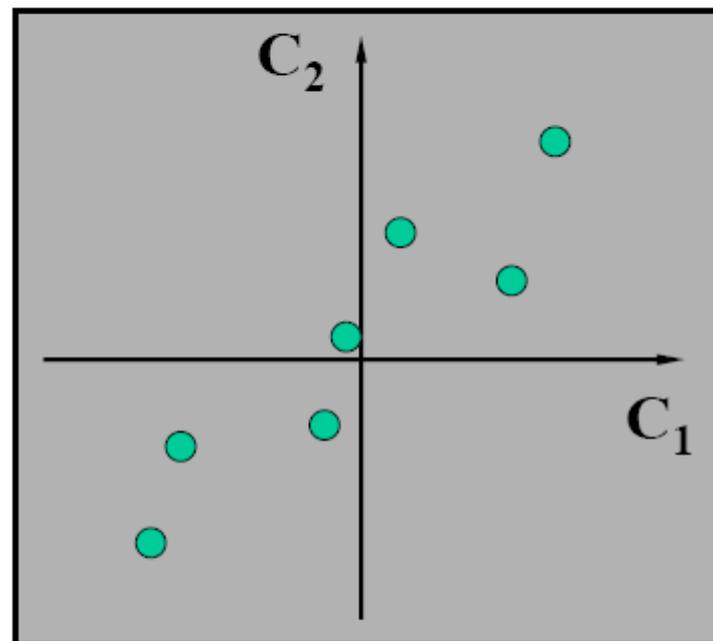
| | X_1 | X_2 |
|-------|-------|-------|
| X_1 | 1 | |
| X_2 | 0.97 | 1 |

Composantes Principales dans le cas de 2 variables

| X_1 | X_2 |
|-------|-------|
| 2.8 | 13.3 |
| 3.9 | 16.6 |
| 8.5 | 17.6 |
| 9.3 | 20.2 |
| 11.2 | 22.1 |
| 14.9 | 23.9 |
| 16.4 | 27.0 |



| C_1 | C_2 |
|-------|-------|
| -6.77 | -6.80 |
| -5.67 | -3.50 |
| -1.07 | -2.50 |
| -0.27 | 0.10 |
| 1.63 | 2.00 |
| 5.33 | 3.80 |
| 6.83 | 6.90 |



Les **données** sont **centrées**
autour de la moyenne

$$c_i = (x_i - \bar{x})$$

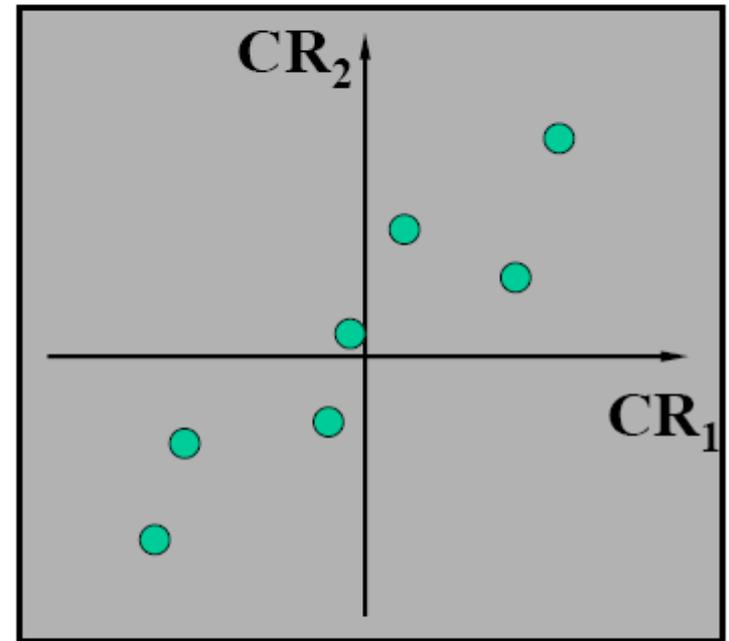


Composantes Principales dans le cas de 2 variables

| X_1 | X_2 |
|-------|-------|
| 2.8 | 13.3 |
| 3.9 | 16.6 |
| 8.5 | 17.6 |
| 9.3 | 20.2 |
| 11.2 | 22.1 |
| 14.9 | 23.9 |
| 16.4 | 27.0 |



| CR_1 | CR_2 |
|--------|--------|
| -1.32 | -1.46 |
| -1.11 | -0.75 |
| -0.21 | -0.54 |
| -0.05 | 0.02 |
| 0.32 | 0.43 |
| 1.04 | 0.81 |
| 1.34 | 1.48 |

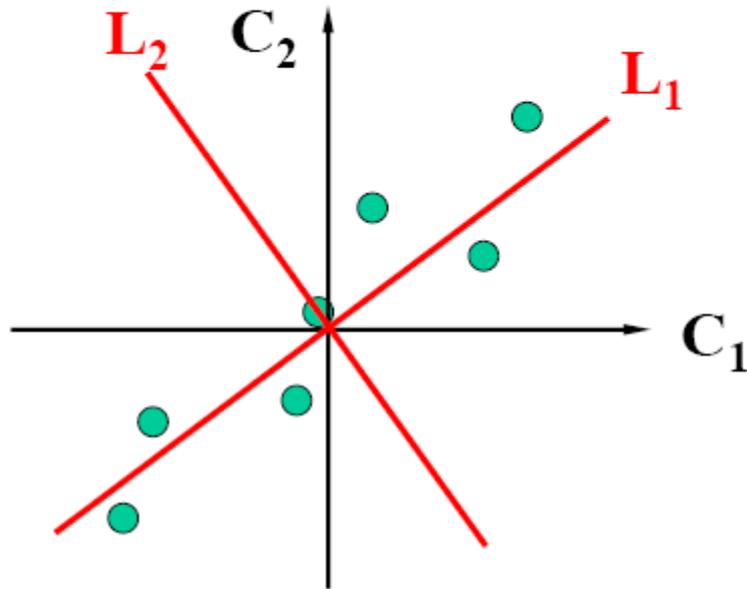


Ou bien, les **données** sont
centrées réduites

$$cr_i = (x_i - \bar{x})/\sigma$$

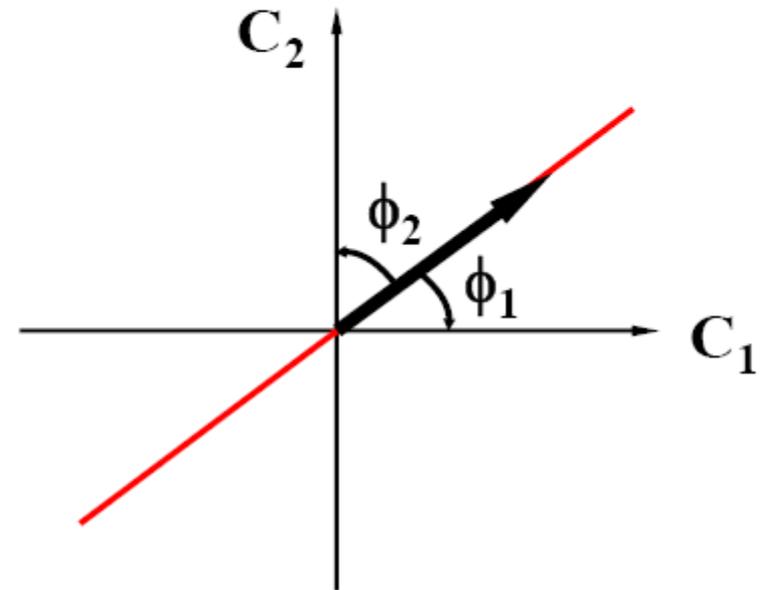


Retenons le passage aux variables **centrées réduites**



La **première composante principale** est la **droite des moindres carrés**.

La **seconde composante** est **orthogonale** à la première.



Projection orthogonale
du vecteur unitaire :

$$p_1 = \cos \phi_1$$

$$p_2 = \cos \phi_2$$

Retenons le passage aux variables **centrées réduites** qui conduit à des nombres sans dimensions :

Ce changement de variable confère **le même poids** à chacun des descripteurs.

Le passage aux variables centrées réduites étant une **transformation linéaire**, il n'y a **pas de changement** de la matrice de corrélation.

Matrice de **corrélation**



| | | |
|-------|-------|-------|
| | X_1 | X_2 |
| X_1 | 1 | |
| X_2 | 0.97 | 1 |

| | | |
|---------------|--------|-------|
| Valeur propre | 1.971 | 0.029 |
| % variance | 98.571 | 1.429 |
| % cumulé | 98.571 | 100 |

Vecteurs propres

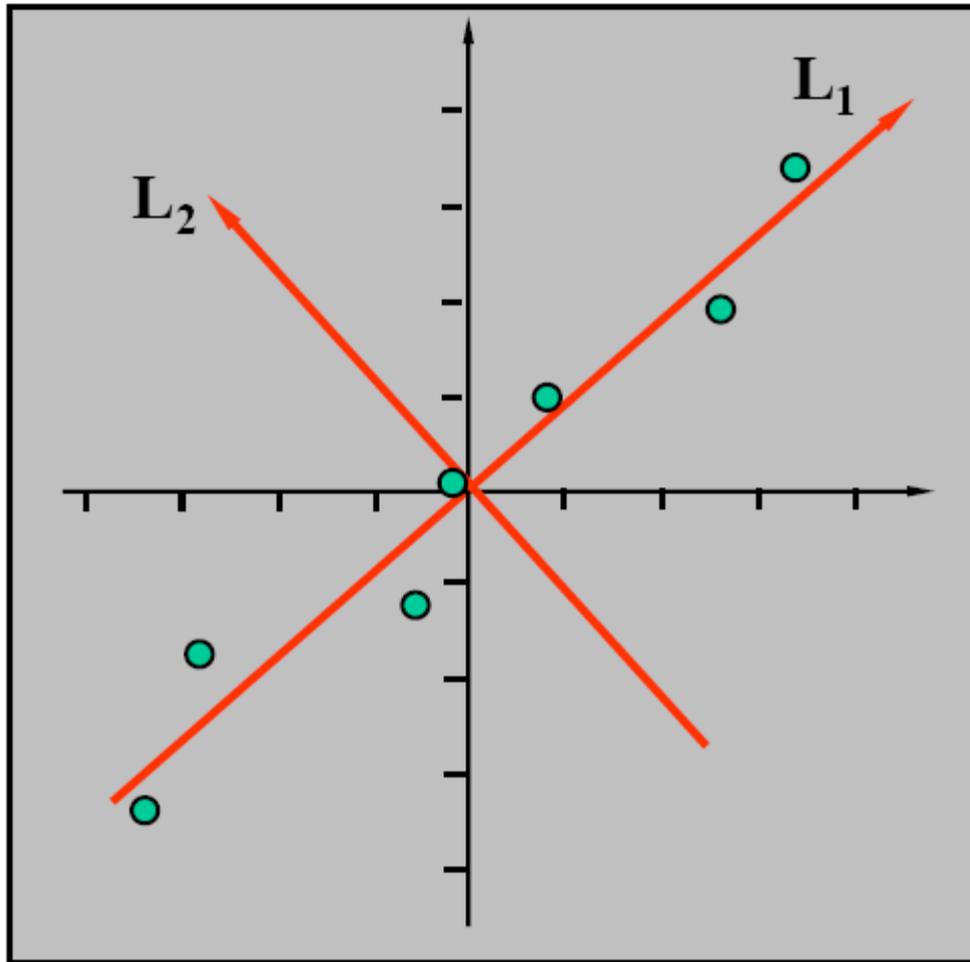
| V1 | V2 |
|-------|--------|
| 0.707 | 0,707 |
| 0.707 | -0,707 |

| V1 | V2 |
|-----|-----|
| a11 | a12 |
| a21 | a22 |

Calcul des coordonnées en composantes principales

$$L1 = CR1 * a11 + CR2 * a21$$

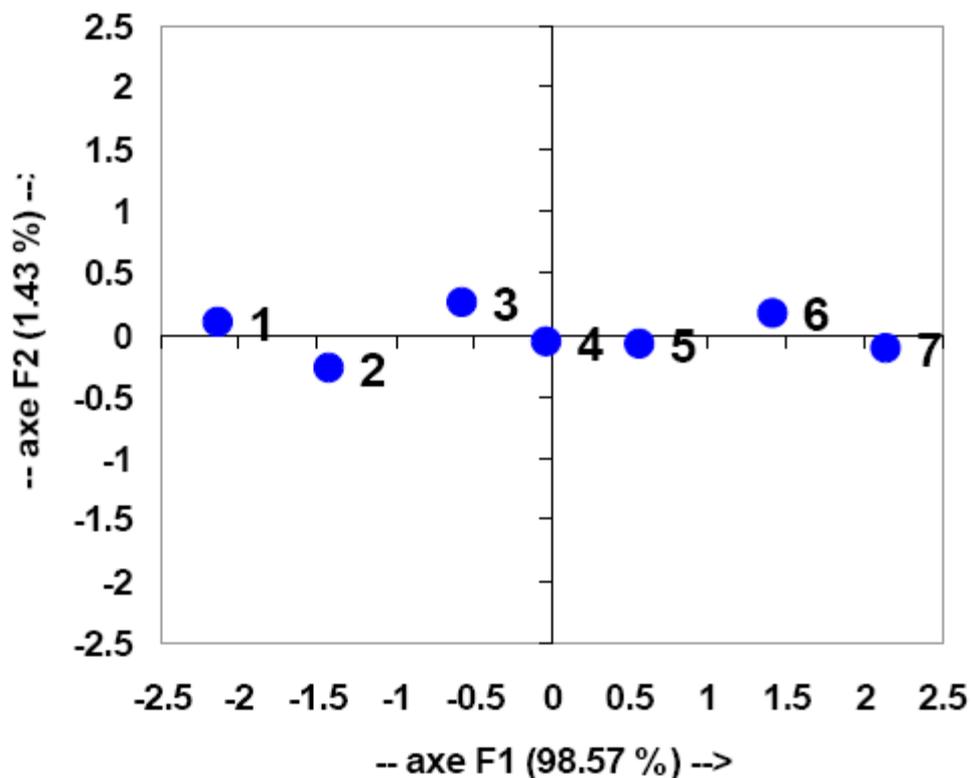
$$L2 = CR1 * a12 + CR2 * a22$$



Les **coordonnées** des individus dans **l'espace des composantes** sont appelées les « **SCORES** ».

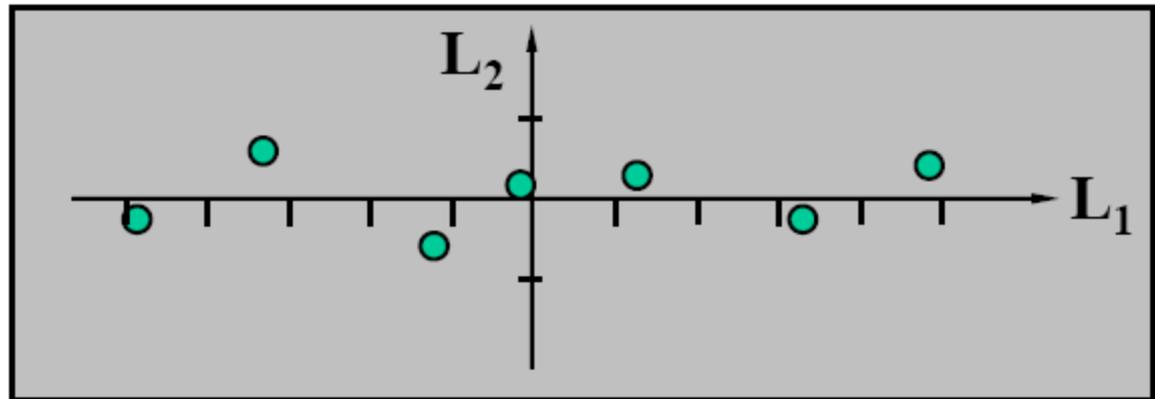
| L_1 | L_2 |
|--------|--------|
| -2.124 | 0.101 |
| -1.420 | -0.274 |
| -0.569 | 0.249 |
| -0.024 | -0.057 |
| 0.570 | -0.084 |
| 1.417 | 0.174 |
| 2.149 | -0.109 |

Individus (axes F1 et F2 : 100.00 %)



| F₁ | F₂ |
|----------------------|----------------------|
| -2.124 | 0.101 |
| -1.420 | -0.274 |
| -0.569 | 0.249 |
| -0.024 | -0.057 |
| 0.570 | -0.084 |
| 1.417 | 0.174 |
| 2.149 | -0.109 |

| L_1 | L_2 |
|--------|--------|
| -2.124 | 0.101 |
| -1.420 | -0.274 |
| -0.569 | 0.249 |
| -0.024 | -0.057 |
| 0.570 | -0.084 |
| 1.417 | 0.174 |
| 2.149 | -0.109 |

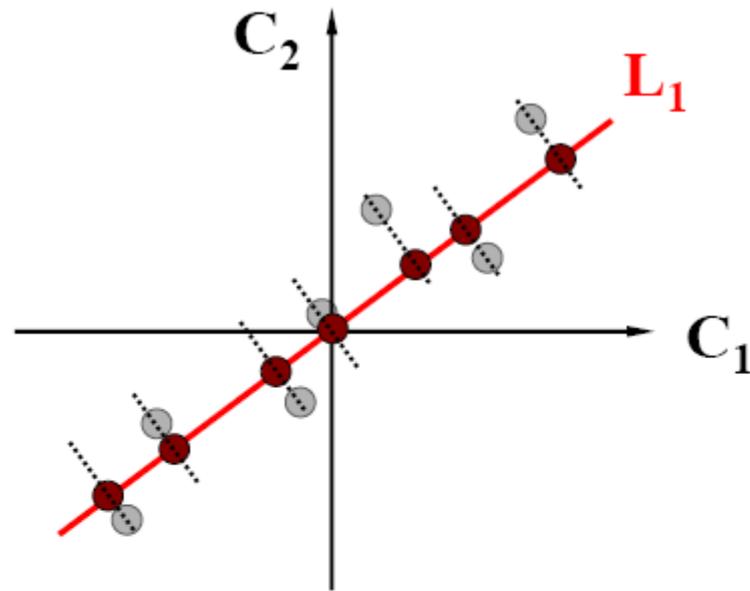


Les descripteurs dans ce nouveau système d'axes sont **totallement d corr l s** :

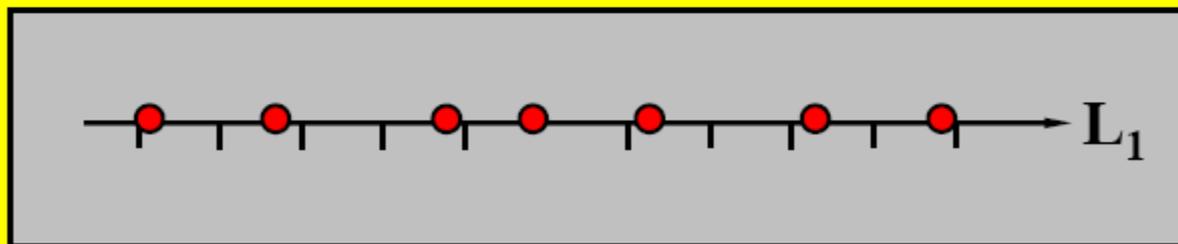


Matrice de
corr lation

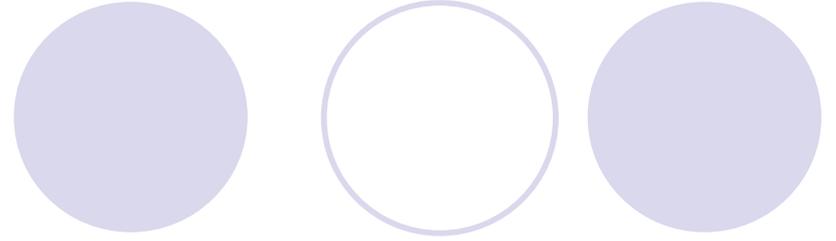
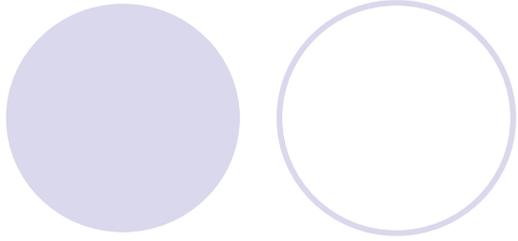
| | L_1 | L_2 |
|-------|-------------|-------|
| L_1 | 1 | |
| L_2 | 0.00 | 1 |

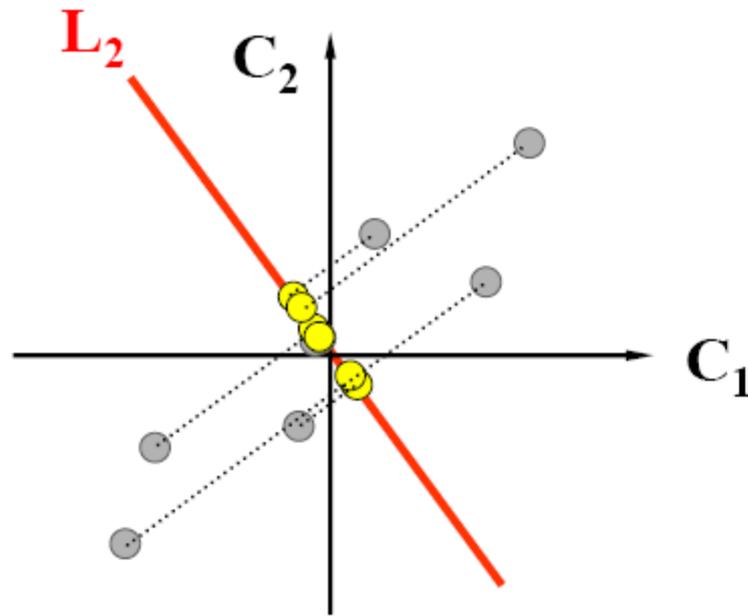


A elles seules, les projections des individus sur la première composante :

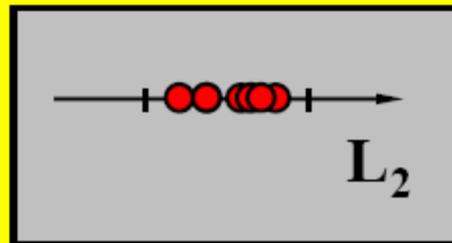


« visualisent ou expliquent » la quasi totalité de la « dispersion » des individus.

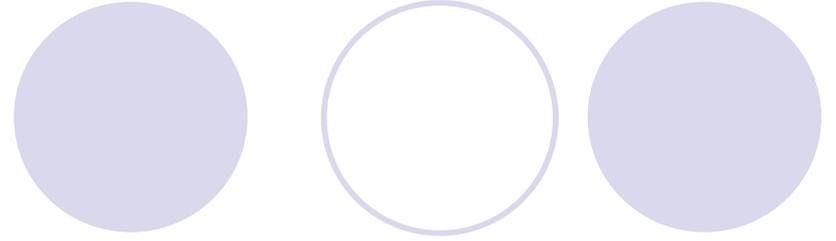
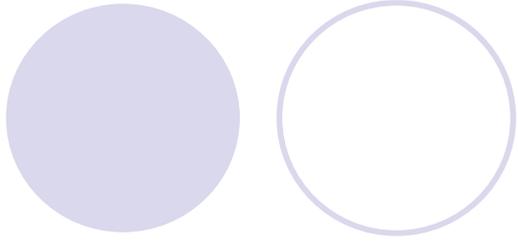


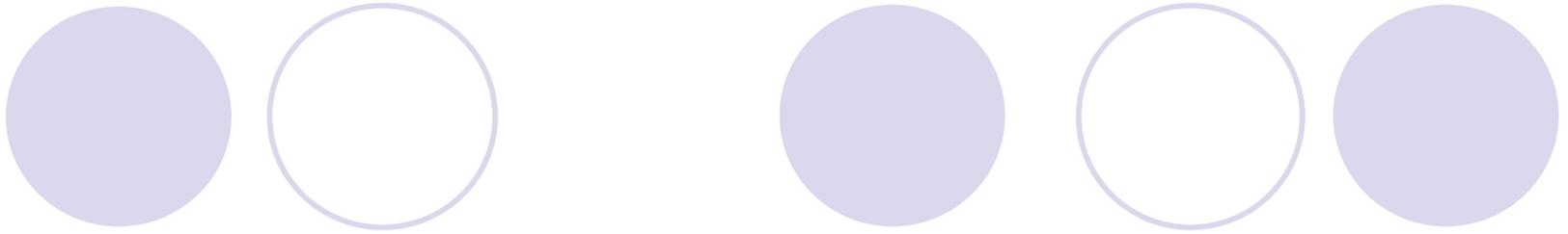


Le fait de négliger la part de « dispersion » expliquée par les projections des individus sur la seconde composante :

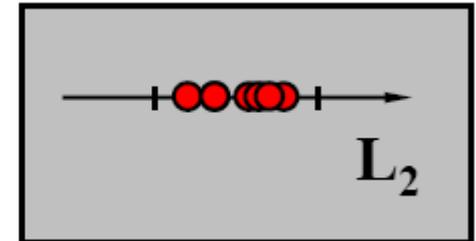
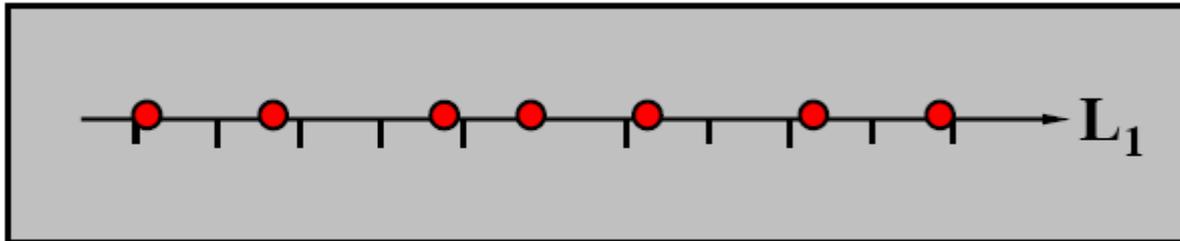


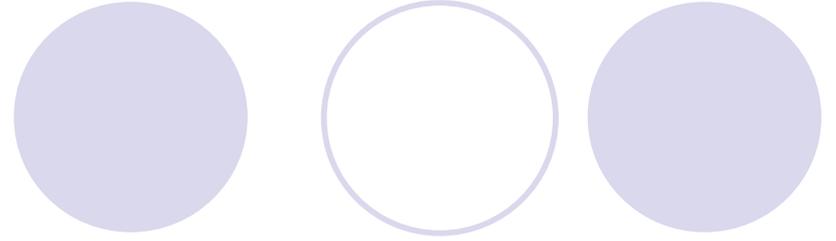
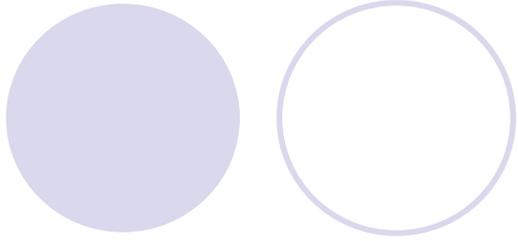
constitue une **perte** d'information **négligeable**.



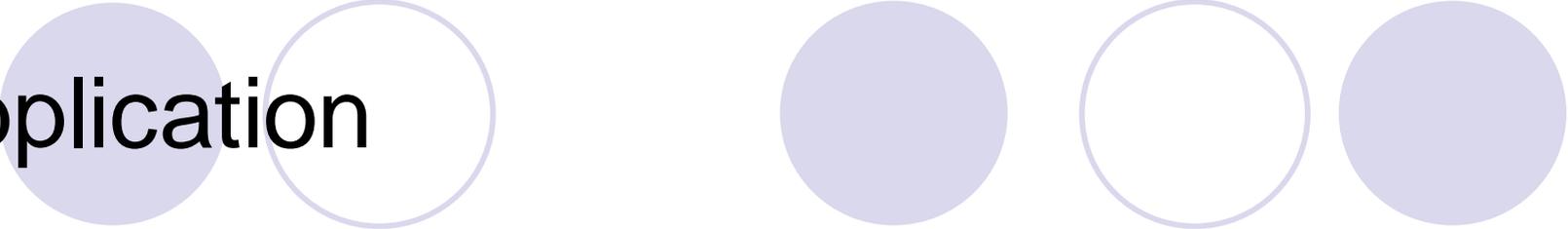


L'information apportée par une projection, dépend fortement de la direction, donc ici de la composante principale considérée ...





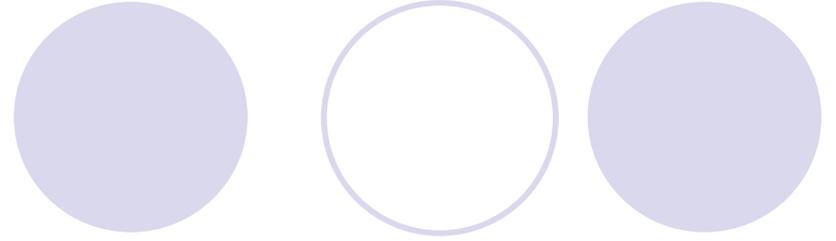
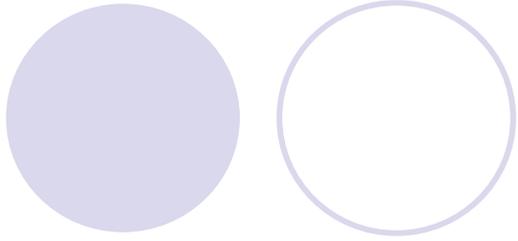
Application



Le levain remplace la levure boulangère utilisée dans les pains industriels et traditionnels. Les levures et les bactéries lactiques du levain font monter la pâte et donnent du goût au pain.

il s'agit d'un mélange de farine , d'eau et de sel qu'on laisse pétrir pendant une vingtaine de minutes dans un pétrin mécanique pour obtenir une pâte élastique qui sera abandonnée une quinzaine de minutes dans le pétrin.

Puis on a mesuré l'**acidité** , **la viscosité**, **la durée et la température** .



Acidité Viscosité Durée Température

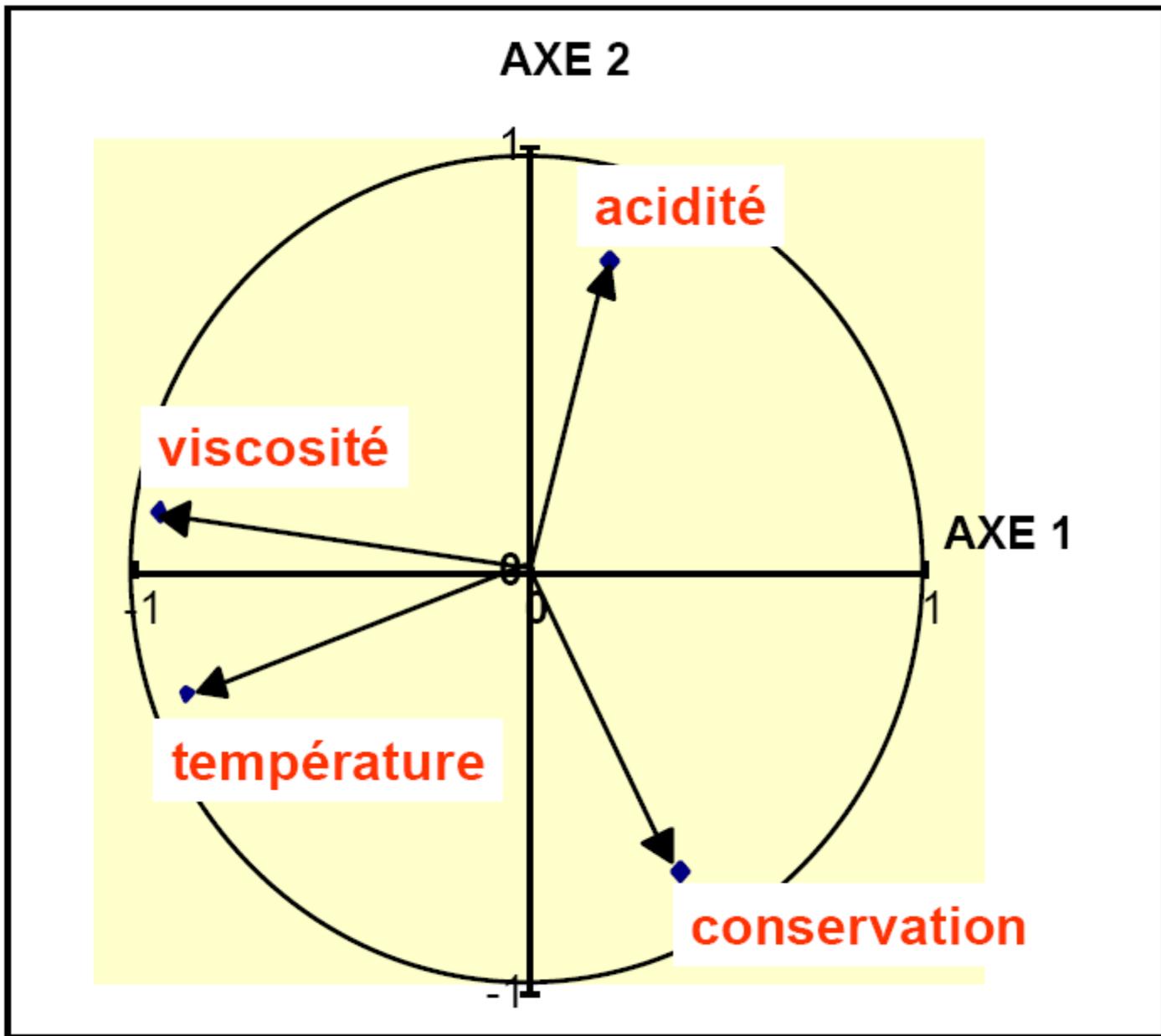
| | | | | | | | |
|------------------|-----------|---------------|--------------|-----------|-----------|-----------|------------------------|
| individus | 1 | 121,5 | 8100 | 1 | 8 | L1 | Type de levains |
| | 2 | 126 | 9360 | 1 | 8 | L2 | |
| | 3 | 124 | 13000 | 1 | 22 | L1 | |
| | 4 | 131,5 | 12260 | 1 | 22 | L2 | |
| | 5 | 147,5 | 6260 | 7 | 8 | L1 | |
| | 6 | 149,5 | 8750 | 7 | 8 | L2 | |
| | 7 | 153,5 | 11250 | 7 | 22 | L1 | |
| | 8 | 154,25 | 12766 | 7 | 22 | L2 | |
| | 9 | 149,5 | 7633 | 15 | 8 | L1 | |
| | 10 | 150,5 | 9366 | 15 | 8 | L2 | |
| | 11 | 155 | 8800 | 15 | 22 | L1 | |
| | 12 | 157,75 | 11100 | 15 | 22 | L2 | |
| | 13 | 152 | 6800 | 21 | 8 | L1 | |
| | 14 | 155 | 9466 | 21 | 8 | L2 | |
| | 15 | 157 | 9400 | 21 | 22 | L1 | |
| | 16 | 159 | 9700 | 21 | 22 | L2 | |

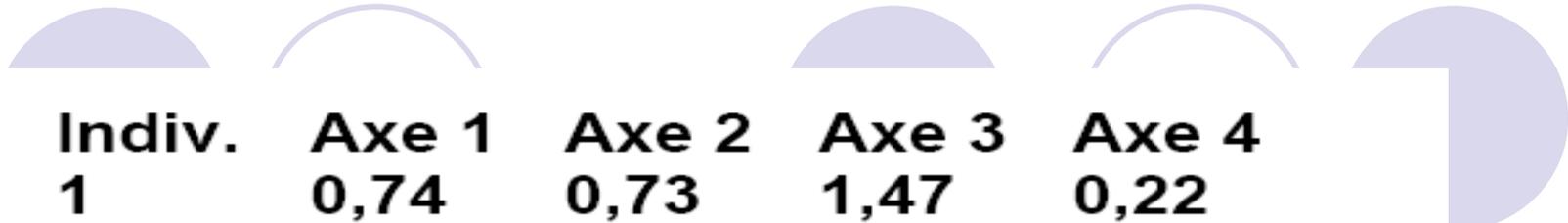
CORRELATIONS DES VARIABLES INITIALES

| | ACIDITE | VISCOSITE | CONSERVATION | TEMPERATURE |
|--------------|---------|-----------|--------------|-------------|
| ACIDITE | 1 | | | |
| VISCOSITE | -0,02 | 1 | | |
| CONSERVATION | -0,072 | -0,341 | 1 | |
| TEMPERATURE | -0,194 | 0,726 | 0 | 1 |

ANALYSE DES VARIABLES SYNTHETIQUES

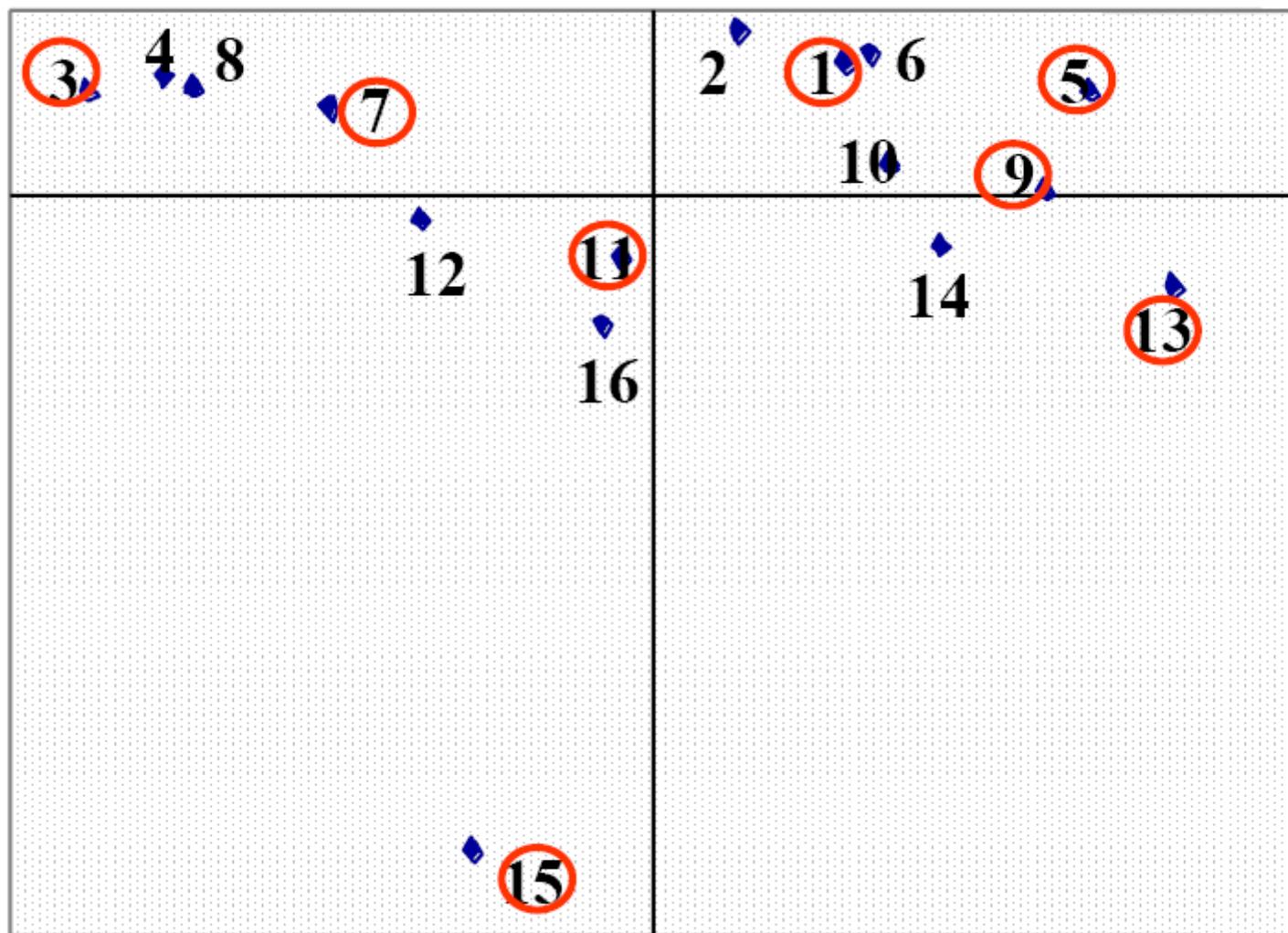
| | AXE 1 | AXE 2 | AXE 3 | AXE 4 |
|--------------|-------|-------|-------|-------|
| ACIDITE | 0,2 | 0,73 | -0,65 | 0,05 |
| VISCOSITE | -0,94 | 0,14 | -0,12 | -0,3 |
| CONSERVATION | 0,38 | -0,71 | -0,58 | -0,12 |
| TEMPERATURE | -0,87 | -0,29 | -0,27 | 0,28 |





| Indiv. | Axe 1 | Axe 2 | Axe 3 | Axe 4 |
|---------------|--------------|--------------|--------------|--------------|
| 1 | 0,74 | 0,73 | 1,47 | 0,22 |
| 2 | 0,32 | 0,89 | 1,31 | -0,2 |
| 3 | -2,2 | 0,57 | 0,54 | -0,22 |
| 4 | -1,91 | 0,66 | 0,45 | -0,06 |
| 5 | 1,69 | 0,57 | 0,62 | 0,72 |
| 6 | 0,84 | 0,77 | 0,43 | -0,14 |
| 7 | -1,27 | 0,48 | -0,38 | 0,26 |
| 8 | -1,79 | 0,59 | -0,49 | -0,27 |
| 9 | 1,51 | 0,03 | -0,14 | -0,04 |
| 10 | 0,91 | 0,16 | -0,27 | -0,64 |
| 11 | -0,13 | -0,33 | -0,89 | 0,82 |
| 12 | -0,91 | -0,13 | -1,09 | 0,03 |
| 13 | 2,02 | -0,49 | -0,62 | 0,04 |
| 14 | 1,11 | -0,26 | -0,84 | -0,88 |
| 15 | -0,71 | -3,55 | 1,42 | -0,06 |
| 16 | -0,21 | -0,7 | -1,51 | 0,3 |

Axe 2



Axe 1