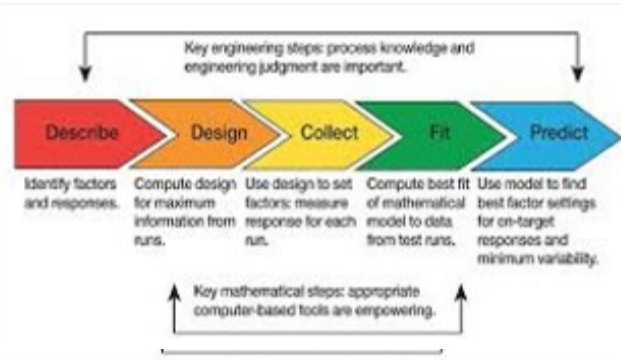


# MTH8301 Planification et analyse d'expériences

## EXERCICES

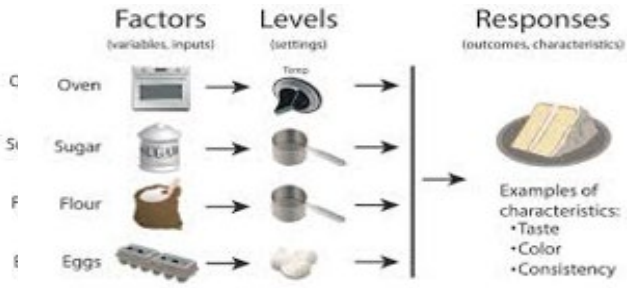
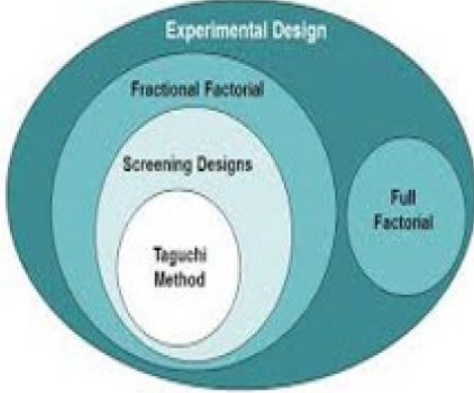


Available Factorial Designs (with Resolution)

Run	2	3	4	5	6	7	8	9	10	11	12	13	14	15
4	Full	Full												
8	Full	IV	III	III	III									
16	Full	IV	IV	IV	IV	III	III	III	III	III	III	III	III	III
32	Full	IV	IV	IV	IV	IV	IV	IV	IV	IV	IV	IV	IV	IV
64	Full	IV	IV	IV	IV	IV	IV	IV	IV	IV	IV	IV	IV	IV
128	Full	IV	IV	IV	IV	IV	IV	IV	IV	IV	IV	IV	IV	IV

Available Resolution III Plackett-Burman Designs

Factors	Runs	Factors	Runs	Factors	Runs
2-7	12, 20, 24, 28, ..., 48	20-23	24, 28, 32, 36, ..., 48	36-39	40, 44, 48
8-11	12, 20, 24, 28, ..., 48	24-27	28, 32, 36, 40, 44, 48	40-43	44, 48
12-15	20, 24, 28, 36, ..., 48	28-31	32, 36, 40, 44, 48	44-47	48
16-19	20, 24, 28, 32, ..., 48	32-35	36, 40, 44, 48		



Nuclear Instruments and Methods in Physics  
 Research Section A: Accelerators, Spectrometers,  
 Detectors and Associated Equipment



Available online 26 November 2022, 167873  
 In Press, Journal Pre-proof

### Toward artificial-intelligence assisted design of experiments

Tommaso Dorigo MODE Collaboration

Les données sont disponibles en fichier Statistica ([STW](https://cours.polymtl.ca/mth6301/WEB-mth8301/MTH8301-Data-Devoirs.stw)) et en fichier Excel ([XLSX](https://cours.polymtl.ca/mth6301/WEB-mth8301/MTH8301-Data-Devoirs.xlsx))  
<https://cours.polymtl.ca/mth6301/WEB-mth8301/MTH8301-Data-Devoirs.stw>  
<https://cours.polymtl.ca/mth6301/WEB-mth8301/MTH8301-Data-Devoirs.xlsx>

No	Nom - Thème	Données	Page
Mth8301-exer-01	erreur expérimentale	-----	3
Mth8301-exer-02	Ishikawa	-----	3
Mth8301-exer-03	fluide	fichier à faire	4
Mth8301-exer-04	matériau	isolant	4
Mth8301-exer-05	éléments expérience	-----	5
Mth8301-exer-06	type d'étude statistique	-----	5
Mth8301-exer-07	électronique	semiconducteur	6
Mth8301-exer-08	métallurgie	powder	7
Mth8301-exer-09	sols contaminés	plomb	8
Mth8301-exer-10	technologie surface	SMT	9
Mth8301-exer-11	expérience avec 3 facteurs-1	-----	10
Mth8301-exer-12	expérience avec 3 facteurs-2	-----	10
Mth8301-exer-13	expérience avec 3 facteurs-3	-----	10
Mth8301-exer-14	expérience avec 3 facteurs-4	-----	11
Mth8301-exer-15	expérience avec 5 facteurs	extraction	12
Mth8301-exer-16	étude en marketing	marketing	13
Mth8301-exer-17	blocage, carré latin	yield	15
Mth8301-exer-18	efficacité médicaments	pain	16
Mth8301-exer-19	stabilité viscosité	food	18
Mth8301-exer-20	hayon voiture	hatchback	22
Mth8301-exer-21	éclairage	candela	23
Mth8301-exer-22	wind tunnel	aircraftWings	24
Mth8301-exer-23	multifactorielle	multifacteurs	26
Mth8301-exer-24	recherche article	de l'article	27
Mth8301-exer-25	conception plan fractionnaire	-----	28
Mth8301-exer-26	plan avec contraintes	-----	28
Mth8301-exer-27			29
Mth8301-exer-28			
Mth8301-exer-29			
Mth8301-exer-30			
Mth8301-exer-31			
Mth8301-exer-32			
Mth8301-exer-33			
Mth8301-exer-34			
Mth8301-exer-35			

### Mth8301-exer-01 : Erreur expérimentale

**No 1a)** Dans la conduite de toute expérience planifiée selon les principes statistiques.

On suppose la présence d'une erreur expérimentale car

- a1) il y a une possibilité de commettre une erreur dans l'exécution d'un essai ;
- a2) les appareils de mesure ne donnent pas toujours des valeurs justes ;
- a3) l'erreur est humaine et il faut toujours se préparer en conséquence ;
- a4) il est difficile de tout contrôler dans une expérience ;
- a5) aucune des réponses ci – haut

Réponse-No1a)

.....

**No1b)** Proposer une définition de l'erreur expérimentale.

Réponse-No1b)

.....

.....

**A)** Proposer une méthode directe qui permet de détecter directement la présence de l'erreur expérimentale d'une variable de réponse Y en présence de plusieurs variables (facteurs) contrôlés

X1, X2, ..., Xk lors de l'exécution d'une série d'essais (tests) en faisant varier les facteurs.

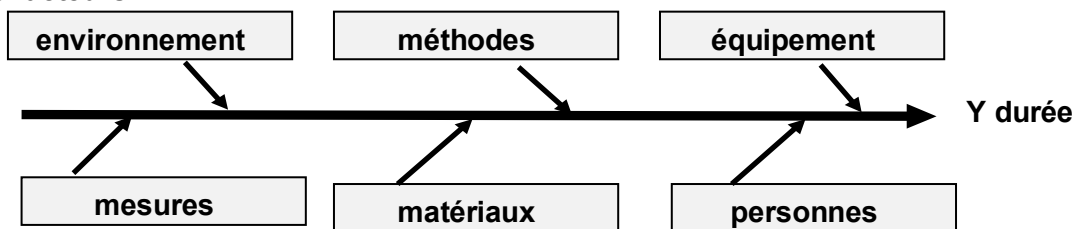
Réponse-No1c) :

.....

.....

### Mth8301-exer-02 : Diagramme de Ishikawa (diagramme de causes à effet)

Construire un diagramme d'Ishikawa décrivant les variables impliquées dans le processus décrivant le trajet entre son domicile et son lieu de travail (ou d'étude). Dresser la liste des facteurs X susceptibles d'affecter la variable de réponse Y mesurant la durée du trajet. Préciser si les facteurs sont contrôlables (on peut fixer leurs valeurs), non contrôlables (facteurs de bruit). Classifier chaque X selon qu'il est catégorique, ou continu. Viser une liste d'au moins 15 facteurs.



#### Remarque

- il n'y a pas une bonne et unique réponse à cet exercice.
- Statistica possède une procédure pour tracer un diagramme d'Ishikawa. Voir *Statistics... Industrial Statistics&6 Sigma... Process analysis... Cause-effect (Ishikawa, Fishbone) diagrams*

### Mth8301-exer-03 : Étude sur quatre 4 types de fluides

Source : D. G. Montgomery, Design and Analysis of Experiments, 7<sup>th</sup> ed., 2008, p. 117

Données = **fluides** à créer dans les filières

MTH8301-Data-devoirs.stw ou MTH8301-Data-devoirs.xlsx

Effectuer une analyse statistique comparative de 4 types de fluides A, B, C D, servant à l'isolation dans les réseaux électriques. La variable de réponse Y est la vie utile (en heures) du fluide. Elle fut mesurée avec des tests accélérés à 35 Kv. Le tableau des résultats des tests est :

X : types-Y : durée en heures						
A	17,6	18,9	16,3	17,4	20,1	21,6
B	16,9	15,3	18,6	17,1	19,5	20,3
C	21,4	23,6	19,4	18,5	20,5	22,3
D	19,3	21,1	16,9	17,5	18,3	19,8

### QUESTIONS

**3a) Les** fluides diffèrent –ils globalement ? Employez un seuil alpha de 0,05.

remarque ; employer un seuil alpha (erreur de première espèce) égale à 0,05 dans tous les exercices

**3b)** Effectuez une analyse des résidus. Les hypothèses de base sont-elles respectées ?

**3c)** Effectuez un test de comparaison multiple (2 à 2) des fluides avec la méthode de Tukey (HSD).

**3d)** Quel fluide devrait-on choisir si on veut maximiser la vie utile?

### Mth8301-exer-04 : Efficacité de matériaux isolants

Source : D. G. Montgomery, Design and Analysis of Experiments, 7<sup>th</sup> ed., 2008, p. 118

Données = **Isolant** disponible dans les filières

MTH8301-Data-devoirs.stw ou MTH8301-Data-devoirs.xlsx

### Problématique

Une expérience fut conduite pour étudier l'efficacité de 5 matériaux isolants.

Des tests à haut voltage furent réalisés afin d'accélérer le temps pour l'obtention d'une panne.

La variable de réponse Y est le temps pour obtenir la première panne (minutes).

Voici le tableau des données de 4 échantillons de chaque matériau.

X : Matériau	Y : temps panne				
1	1	2	4	18	
2	2	5	7	29	
3	110	157	178	194	
4	880	1256	4355	5276	
5	495	7040	5307	10050	

### QUESTIONS

**4a)** Les 5 matériaux ont-ils le même effet sur le temps moyen de la panne ?

**4b)** Faire une analyse des résidus.

**4c)** Faire le graphique du log (moyenne) vs log (écart type).

**4d)** Transformer la variable de réponse et refaire une autre analyse pour comparer les matériaux.

**4e)** Comparer les matériaux à l'aide du test non paramétrique Kruskal-Wallis.

## Mth8301-exer-05 : Composantes d'une expérience

Une expérience est caractérisée par :

- des **traitements** (procédures) définis par les combinaisons des modalités (niveaux) de un ou plusieurs facteurs X contrôlés que l'on veut comparer
- des **unités expérimentales** (UE), idéalement les plus homogènes possibles
- la méthode d'assignation (**protocole**) des traitements aux UE
- les **réponses Y** mesurées

### PROBLÈME

On veut étudier (comparer) l'effet de 2 diètes sur la croissance des enfants.

#### Description des diètes

dA : menus équilibrés de protéines, fruits, légumes avec beaucoup de brocoli

dB : riche en croustilles (chips) et barres de chocolat

On veut déterminer la diète qui aide les enfants à grandir le plus vite possible.

On dispose d'un groupe de 20 enfants et on considère 5 méthodes (protocoles) pour assigner les enfants aux diètes.

#### Description des protocoles

p1 : les laisser choisir

p2 : assigner dA pour les 10 premiers enfants dA et assigner dB pour les 10 autres

p3 : alternance : dA, dB, dA, dB,...

p4 : pour chaque enfant, lancer une pièce de monnaie :

assigner l'enfant à dA si le résultat est pile, assigner l'enfant à dB si le résultat est face

p5 : choisir 10 enfants au hasard et les assigner à dA, les 10 autres sont assignés à dB

### QUESTIONS

5a) Identifier toutes les composantes de l'expérience

5b) Commenter chacun des éléments de 5a)

5c) Décrire les bénéfiques (avantages) et les risques (inconvenients) des protocoles

## Mth8301-exer-06 : Étude statistique : observationnelle ou étude expérimentale ?

Classer chacune des quatre études suivantes selon qu'elle est observationnelle ou expérimentale. Pour chaque étude:

- identifier les variables X (explicatives) les variables Y (réponse),
- préciser la nature catégorique ou continue des variables,
- identifier unités statistiques sur lesquelles les données sont mesurés ou observées,

### QUESTIONS

6a) On compare 2 algorithmes selon le temps de CPU pour réaliser les mêmes 6 problèmes tests.

6b) Un psychologue mesure le temps de réponse de 20 sujets (10 hommes, 10 femmes) soumis à 2 stimuli qui sont administrés dans un ordre aléatoire pour chaque sujet.

6c) Un groupe de fumeurs et un autre groupe de non-fumeurs sont examinés durant une période de 10 ans afin de déterminer s'ils développent une forme de cancer.

6d) Une agence de marketing a développé 2 publicités différentes pour un produit détergent. Afin de comparer l'efficacité des 2 publicités, un prétest est conduit sur un échantillon de 50 personnes est divisé au hasard en 2 groupes. Chaque groupe visualise une des 2 publicités et évalue avec un score de 1 à 10 selon l'échelle (1 = pas efficace) à (10 = très efficace)

## Mth8301-exer-07 : Industrie électronique des semi-conducteurs

Données = **Semiconducteur** disponible dans les filières

MTH8301-Data-devoirs.stw ou MTH8301-Data-devoirs.xlsx

L'expérience a pour but d'augmenter le rendement de fabrication de semi-conducteurs.

Plan factoriel  $2^5$  sans répétition + 4 essais au centre.

Facteurs et leurs modalités

codage / modalités

LETTRE	NOM	- 1	+ 1
A	aperture setting (catégorique)	small	large
B	temps exposition (continu)	- 20% en dessous valeur nominale	+ 20% au-dessous valeur nominale
C	temps development (seconde)	30	45
D	dimension masque (catégorique)	small	large
E	etch time (minute) (continu)	14,5	15,5

La variable de réponse est  $Y = \text{Yield (lbs)}$ . On vise le maximum.

### QUESTIONS

#### **7a) Analyse 1 : exclure les essais au centre (filtre actif)**

##### Éléments à présenter

- Graphique des effets sur échelle demi normale et Pareto
- Tableau d'analyse de la variance
- Tableau des effets et tests de signification
- Analyse des résidus
- Proposer un modèle de prédiction avec seulement les effets significatifs retenus; ce modèle sera employé en 7b)
- Représenter graphiquement la réponse : effets principaux et d'interaction du modèle proposé
- Déterminer les conditions optimales d'opération.

#### **7b) Analyse 2 : inclure les essais au centre (filtre inactif)**

Refaire l'analyse en 6a) avec le modèle retenu en 6a) et en incluant un test de manque d'ajustement.

#### **7c) Serait-il approprié de faire des essais additionnels? Si oui, lesquels?**

## Mth8301-exer-08 : Industrie métallurgique

Données = **Powder** disponible dans les filières

MTH8301-Data-devoirs.stw ou MTH8301-Data-devoirs.xlsx

Expérience pour améliorer le procédé de fabrication de poudre métallique à base d'argent.

Ce produit est employé dans un grand nombre de produits industriels.

### Facteurs

A : Ammonium (%)	modalités : 2 et 30
B : Stir Rate (RPM)	modalités : 100 et 150
C : Temperature (deg. C)	modalités : 8 et 40

### Réponses

Y1\_den : density (g/cm<sup>2</sup>)      Y2\_sur : surface area (cm<sup>2</sup>/g)

Plan complet 2<sup>3</sup> avec une répétition (n = 2)

### Objectifs d'optimisation

Maximiser Y1\_den : au moins 14 (intervalle satisfaisant : 14 et plus)

Cibler Y2\_sur : cible = 0,45 (intervalle satisfaisant : 0,3 à 0,6)

## QUESTIONS

**8a)** Analyser la réponse Y1-den

### Éléments à considérer

Estimation des coefficients d'un modèle approprié, présentation du tableau ANOVA

Graphique Pareto des effets, identification des effets significatifs, graphiques de la réponse

**8b)** Effectuer une analyse diagnostique des résidus. Commenter les résultats.

Refaire l'analyse avec la transformation inverse  $Z1 = 1 / Y1\_den$

Les problèmes furent-ils corrigés ?

**8c)** Construire des courbes contours pour visualiser et interpréter la réponse Y1\_den et son objectif.

**8d)** Exécuter une analyse similaire comme (8a)-(8b)-(8c) pour la réponse Y2\_sur

Remarque : Y2\_sur n'a pas besoin d'être transformée

**8e)** Faire l'optimisation simultanée de Y1\_den et Y2\_sur.

Déterminer les conditions d'opération qui permettent d'atteindre les exigences désirées en utilisant les fonctions de désirabilité.

Représenter graphiquement et conjointement sur un même graphique (overlaid)

les 2 réponses Y1\_den et Y2\_sur.

Utiliser la procédure graphique de Statistica :

*Graphs ...3DXYZ Plots...Contour Plots...Overlaid Contour Plots*

**8f)** On vous informe après avoir présenté vos analyses que les essais furent réalisés par 2 opérateurs. Les essais de la répétition no 1 furent faits par l'opérateur OP1 et les essais de la répétitions no 2 furent faits par l'opérateur OP2.

Refaire l'analyse de la variable Z1 pour tenir compte de cette nouvelle information.

Présenter une conclusion sommaire (quelques phrases).

## Mth8301-exer-09 : remédiation de sols contaminés

Données = **Plomb** disponible dans les filières  
MTH8301-Data-devoirs.stw ou MTH8301-Data-devoirs.xlsx

**Problématique** : expérience en laboratoire pour comparer l'efficacité de deux matériaux.

**But** comparaison de 2 matériaux Fisbone et Hydroxyapatite  
utilisation de Fisbone (moins cher) comme un substitut de Hydroxyapatite  
pour remédier des milieux aqueux contaminés au plomb (Pb)

**Plan** complet  $2^3$  avec une répétition ( $n = 2$ )

### Facteurs

A_Apatite	contaminant Apatite (mg) modalités = 30 et 60 (ajout à de l'eau dé-ionisée)
B_pH	modalités = 5 et 7
C_Pb	concentration plomb (nM ou ppm) modalités = 0,48 (= 100 ppm) et 2,41 (= 500 ppm)

### Réponses

Y1_Pb_F	: concentration (nM) plomb (Pb) avec Fishbone (F)
Y2_Pb_H	: concentration plomb (Pb) avec Hydroxyapatite (H)
Y3_pH_F	: pH avec Fishbone (F)
Y4_pH_H	: pH avec Hydroxyapatite (H)

## Questions

**9a)** Analyser la réponse Y1\_Pb\_F

### Éléments d'analyse à présenter

estimation des coefficients d'un modèle approprié, présentation du tableau ANOVA, identification des effets significatifs, analyse Pareto, production de graphiques de la réponse ; analyse diagnostique des résidus.

**9b)** Analyser la réponse Y2\_Pb\_H éléments à considérer : comme en **9a)**

**9c)** Résumer les 2 analyses précédentes sous la forme d'un tableau.

Présenter une conclusion sommaire sur la comparaison entre les 2 matériaux.

**9d)** Les auteurs de l'article sur lesquels sont basés ces données concluent que

Le matériau Fishbone est aussi bon que le Hydroxyapatite.

Ils recommandent son adoption.

Êtes-vous d'accord avec cette conclusion ? Justifier votre réponse.

**9e)** Proposer une autre manière de structurer les données enfin de comparer l'efficacité des 2 matériaux à réduire le plomb via la variable  $Y = \text{concentration Pb}$ .

Présenter un tableau montant cette nouvelle structure des données.

Exécuter une nouvelle analyse des données avec cette structure afin de comparer l'efficacité des 2 matériaux. Comparer la conclusion de cette analyse avec celle de **9d)**



## Mth8301-exer-10 : Soldering and Surface Mount Technology

Données = **SMT** disponible dans les filières  
MTH8301-Data-devoirs.stw ou MTH8301-Data-devoirs.xlsx

Plan fractionnaire  $2^{8-3}$  aucune répétition (n = 1)

Facteurs	modalités	
1. A: Squeegee Pressure (MPA)	0,1	0,3
2. B: Printing Speed (mm/s)	24	32
3. C: SqueegeeAngle (deg)	45	65
4. D: Temperature (deg C)	20	28
5. E: Viscosity (kCps*1000)	1125	1275
6. F: Cleaning Interval (stroke)	8	15
7. G: SeparationSpeed (mm/s)	0,4	0,8
8. H: Relative Humidity (%)	30	70

Réponses Y1\_PVM Pourcentage Volume Matching  
Y2\_NPU Non-conformités Per Unit

Objectifs cibler Y1\_PVM cible = 1 intervalle acceptable : 0,95 à 1,05  
minimiser Y2\_NPU cible = 0 intervalle acceptable : 0 à 10

### QUESTIONS

- 10a)** Déterminer comment furent définis les générateurs de cde plan fractionnaires
- 10b)** Préciser les effets d'interaction (d'ordre 2 seulement) qui sont confondus
- 10c)** Analyser chacune des 2 réponses avec un modèle approprié.  
Présenter une conclusion sommaire.  
Ne pas inclure l'analyse diagnostique de résidus.

**Définition** : un facteur X est déclaré *inerte* par rapport à une variable de réponse Y s'il est *ni significatif ou ni important* suite à analyse statistique.

La déclaration qu'un facteur X est inerte n'est pas générale mais conditionnelle selon

- l'espace de variation auquel X fut assujetti au cours des essais
- la variable de réponse Y.

- 10d)** Compte tenu du résultat en **10c)**, expliquer pourquoi les expériences avec un grand (disons 6 ou plus) nombre de facteurs ne supporte pas fortement le principe de la répétition des essais.  
Expliquer le rôle caché des facteurs inertes.
- 10e)** Déterminer des conditions optimales d'opération pour opérer le procédé.

### **Mth8301-exer-11 : Expérience avec trois facteurs--Partie 1**

Un plan factoriel complet avec trois facteurs A, B, C chacun variant à deux modalités (- et +) a été réalisé pour étudier leurs effets sur une variable de réponse Y.  
Le plan a été répété une autre fois ( $n = 2$ ).

#### **QUESTIONS justifier vos réponses**

- 11a)** Combien de traitements distincts sont présents dans cette expérience ?
- 11b)** Combien d'essais furent réalisés ?
- 11c)** Préciser l'avantage et l'inconvénient d'avoir répété le plan.
- 11d)** Quelle serait votre stratégie pour le choix des modalités de chaque facteur ?
- 11e)** Dressez la liste de tous les effets que l'on peut estimer.
- 11f)** Donnez le nom d'une méthode statistique graphique avec laquelle on peut visualiser les résultats des essais ;

### **Mth8301-exer-12 : Expérience avec trois facteurs--Partie 2 – suite de Mth-exer-11**

- 12a)** Précisez les critères de votre méthode qui permet de s'assurer que l'expérience a été une *réussite*. Précisez le sens de ce que vous entendez par *réussite*.
- 12b)** Donnez le nom de la méthode statistique principale avec laquelle les résultats observés de Y seront analysés. Comment s'appelle la synthèse numérique de cette analyse ?
- 12c)** Présenter les principaux éléments de cette analyse.
- 12d)** L'erreur expérimentale peut-elle être estimée directement sans ajuster de modèle? Donnez la formule pour le calcul de cette estimation et les degrés de liberté de cette estimation.
- 12e)** La méthode statistique proposée en 12b) permet-elle d'estimer l'erreur expérimentale ?
- 12f)** Peut-on tester si le modèle est adéquat pour représenter le processus indépendamment de la valeur du  $R^2$  et du  $R^2$  ajusté ? Expliquer.

### **Mth8301-exer-13 : Expérience avec trois facteurs--Partie 3 - suite de Mth-exer-11**

Supposons que le test F montre que l'effet d'interaction de A et B est significative.

#### **QUESTIONS**

- 13a)** Expliquez en mots et illustrez graphiquement.  
Supposons que l'analyse statistique montre que les facteurs A, B, C n'ont aucun effet principaux et d'interaction statistiquement significatif sur Y dans l'espace expérimental
- 13b)** Décrire le comportement du graphique de Y
- 13c)** Proposer une équation de prédiction pour la réponse Y
- 13d)** Proposer un objectif possible de l'expérience qui permettrait de conclure que l'expérience a été un *échec*.
- 13e)** Proposez un objectif possible de l'expérience qui permettrait de conclure que l'expérience a été un *succès*.

**Mth8301-exer-14 : Expérience avec trois facteurs--Partie 4- suite de Mth-exer-11**

Supposons que l'analyse statistique montre

- la moyenne générale de tous les essais est de 100
- l'erreur expérimentale estimée ( $\sigma$ ) est de 2
- le facteur A possède un effet principal significatif de -20
- le facteur B possède un effet principal significatif de 10
- l'effet d'interaction de A et B est significatif et vaut 15
- le facteur C est inerte:
  - son effet principal vaut 1 et est non significatif
  - il n'a pas d'effet d'interaction significatif avec A
  - il n'a pas d'effet d'interaction significatif avec B.

**QUESTIONS**

**14a)** Écrire l'équation de prédiction de Y en fonction des facteurs A B C.

**14b)** Comment pourrait-on décrire un rôle statistique joué par le facteur C ?

**14c)** Illustrez graphiquement l'allure de la réponse Y en fonction des facteurs A et B en plaçant le facteur A sur l'axe horizontal.

**14d)** Si on veut maximiser Y quel serait votre choix pour A et B et C?

Quel serait votre la valeur de la prédiction de Y ?

Quelle serait une recommandation pour le choix de C et précisez votre critère.

**14e)** Si on veut contrôler Y a une valeur cible de 80, quelle serait votre recommandation pour les modalités des facteurs A, B, C ?

## Mth8301-exer-15 : procédé d'extraction de zinc et cuivre

Données = **Extraction** disponible dans les filières  
MTH8301-Data-devoirs.stw ou MTH8301-Data-devoirs.xlsx

### Problématique

Un procédé chimique permet d'extraire des métaux lourds comme le Zinc et le Cuivre. On veut mesurer l'efficacité du procédé en calculant le % d'extraction. On a fait varier 5 facteurs choisis du procédé X1, X2, X3, X4, X5 à 2 modalités avec un plan complet de  $2^{**5}$  (32) de essais. On a ajouté 6 essais au centre pour examiner l'influence potentielle d'effets quadratiques avec les facteurs continus X1 X2 X3 X5. Le facteur X4 étant catégorique on a fait 3 essais au centre avec le sulfate (X4 = -1) et 3 essais avec le chlore (X4 = 1). Pour chaque essai on a mesuré deux variables de réponse

Y\_Zn % d'enlèvement du Zinc    Y\_Cu % d'enlèvement du cuivre

### FACTEURS : nom, type, unité mesure, modalités

nom	type	unité	modalité1	modalité2	modalité3	
			base (-1)	élevée (+1)	centre (0)	
X1	pH solution initiale	continu	pH	4,5	6,5	5,5
X2	concen. initiale du Zn ou Cu	continu	mg/l	25	75	50
X3	concen. D2EHPA	continu	vol	5	10	7,5
X4	solut. aqueuse initiale	catégorique	---	sulfate	chlore	---
X5	taux d'agitation	continu	ram	400	500	450

### QUESTIONS

**15a)** On considère seulement les 32 essais (plan  $2^{**5}$ ) en excluant les 6 essais au centre.

Le fichier contient un filtre pour sélectionner les 32 premières observations.

Voir .... *Tools ... Selection Conditions ... Edit*

Analyser chaque variable de réponse avec un modèle contenant des effets principaux, des interactions doubles et des interaction triples.

Présenter tous les résultats usuels (tableau des effets significatifs, ANOVA, graphique de Pareto, R<sup>2</sup>, R<sup>2</sup> ajustés etc) dans un tableau comparatif des 2 variables de réponse.

Ne pas faire d'analyse des résidus.

**15b)** Pour chaque variable de réponse, refaire une nouvelle analyse avec seulement les effets significatifs retenus dans l'analyse 15a).

**15c)** Avec les modèles retenus en 15b) déterminer les conditions pour maximiser les 2 réponses.

**15d)** Refaire une analyse séparée pour chacune des 2 solutions (sulfate, chlore).  
(colonnes 16 à 31)

Dans ces 2 analyses on a un Plan factoriel complet  $2^{**4}$  et 3 essais au centre.

On utilisera un modèle avec des effets principaux, des effets d'interaction et des effets quadratiques. Tester si le modèle avec des effets quadratiques est nécessaire.

**15e)** Laquelle des 2 approches (analyse globale, analyse séparée) est préférable?  
Justifier votre réponse.

## Mth8301-exer-16 : Étude de marketing : 19 facteurs en 20 essais

Données = **marketing.sta** et **marketing\_corr.sta** disponible dans les filières  
MTH8301-Data-devoirs.stw ou MTH8301-Data-devoirs.xlsx

**Source** : International Journal of Research in Marketing, 2006, vol 23, pp 309-319  
19 facteurs à 2 modalités / plan saturé de Plackett-Burman de 20 essais de résolution III

<u>ID</u>	<u>Facteur Nom</u>	<u>Facteur lettre</u>	<u>(-1) control</u>	<u>(+1) nouvelle idée</u>
1	Envelope Teaser	A_EnvTeaser	General offer	Product-specific offer
2	Return Address	B_Return	AddBlind	Add company name
3	Official Ink-stamp	C_Stamp	yes	no
4	Postage	D_Postage	pre-printed	stamp
5	Additional Graphics	E_AddGraph	yes	no
6	Price Graphic	F_PriceGraph	small	large
7	Sticker	G_Sticker	yes	no
8	PersonalizeLetter	H_PerLetter	no	yes
9	Copy Message	I_CopyMes	targeted	generic
10	Letter Headline	J_LetHead	headline1	headline2
11	List of Benefits	K_Benefits	standard layout	creative layout
12	Postscript on Letter	L_Postscript	std version	new P.S.
13	Signature	M_Signature	manager	senior executive
14	Product Selection	N_Selection	product info	few
15	Value of Free Gift	O_FreeGift	many	low
16	Reply Envelope	P_RepEnvelope	high	new style
17	Information Buckslip	Q_InfoBuckslip	control	free gift info
18	2nd Buckslip	R_2ndBuckslip	product info	yes
19	Interest Rate	S_IntRate	low	high

**Variables de réponse** : Y\_orders et Y\_responRate (%)

Seule la variable Y\_ResponRate sera analysée dans cet exercice.

### QUESTIONS

- 16a)** Montrez que le design de Plackett-Burman est orthogonal.
- 16b)** Montrez que la corrélation entre les effets principaux et les effets d'interaction a une structure particulièrement simple. Précisez cette structure et formuler une conclusion.  
**aide** : *Basic Statistics/Tables .... correlation matrices ....*  
Identifier un sous ensemble de facteurs dominants pour construire un modèle pour  
Interpréter les données.
- 16c)** Ajuster un premier modèle M1 pour Y\_ResponRate avec seulement les effets principaux.  
Identifier les 5 facteurs les plus importants.
- 16d)** Ajuster un deuxième modèle M2 contenant tous les effets principaux et les effets d'interaction.
- 16e)** On peut montrer que le plan de Plackett-Burman de 20 essais est un plan factoriel complet pour 3 facteurs. À la suite de 16c) choisir 3 facteurs et ajuster un modèle M3 contenant des effets principaux et des effets d'interaction.  
Conclure sur la méthode d'une campagne efficace en marketing.

## Aide No16

---

### Plans de Plackett-Burman

Les plans de Plackett-Burman sont généralement non recommandés car ils sont de résolution III. Toutefois la structure de corrélation est relativement simple et peut être exploitée comme cela sera fait dans cet exercice. Le **nombre d'essais est un multiple de 4** : 4 essais, 8 essais, 12 essais, 16 essais, 20 essais, 24 essais, .... ; certains plans coïncident avec les plans de  $2^k$  (8-16-32-64...) Essais. La résolution des plans est de III si le nombre de facteurs est égal au nombre d'essais moins 1 comme c'est le cas dans cet exercice. La matrice de corrélation a une structure simple. Elle est donnée dans l'exercice. Moyennant certaines hypothèses, on peut dégager un modèle basé sur un nombre de facteurs réduits. C'est ce qui sera fait dans cet exercice.

### *Industrial Statistics & six-Sigma .... $2^{k-p}$ max unconfounded or min aberration designs*

Compte tenu de **16b) et 16c)** les 5 effets principaux les plus importants sont partiellement confondus avec des interactions. Un des facteurs identifiés comme important est partiellement confondu avec un effet d'interaction avec une corrélation élevée.

Identifiez ce cas. Proposer une conclusion concernant l'influence réelle véritable de ce facteur.

**16a)** l'orthogonalité du plan Plackett-Burman est démontrée avec la matrice de corrélation

**16b)** Il faut examiner la matrice de corrélation des 19 effets principaux A B C ... S et 171 effets d'interaction AB AC ... RS . Cela nécessite d'examiner une matrice corrélation de 190X 190. Cela est particulièrement encombrant.

Comme alternative on peut y arriver par morceaux.

On peut examiner la structure de corrélation entre les 19 facteurs A B ... S et une série de 18 effets d'interaction, comme AB AC ...AS.

L'examen de cette structure permettra de dégager une structure de corrélation relativement simple. On peut valider que la conclusion précédente se répète avec une autre série d'effets d'interaction, comme BC, BD, ..., BS.

**16c)** Postuler un premier modèle avec tous effets principaux seulement.

Dégager les effets les plus importants.

Refaire l'analyse avec un deuxième modèle contenant les effets principaux importants identifiés dans le premier modèle et en ajoutant dans ce deuxième modèle des effets d'interaction. Identifier les facteurs les plus importants. Il y en a peu.

**16d)** Dégager un modèle définitif à la suite de l'analyse en **16c)**

**16e)** Optimiser la réponse avec le modèle définitif de **16d)**

## Mth8301-exer-17 : Blocage, Carré Latin, Carré Gréco-Latin

Données = **Yield.sta** disponible dans les filières

MTH8301-Data-devoirs.stw ou MTH8301-Data-devoirs.xlsx

### Problématique : Rendement d'une réaction chimique (Yield of a chemical process)

Réponse **Y = Yield** facteur principal **X = Catalyst**

3 autres facteurs : Batch de matériau, concentration Acid, Time

### Facteurs

Batch : 1 à 5 / Acid = 1 à 5 / Time = A B C D E / Catalyst = a b c d e

Time est recodée Time2 : A=1 B=2 C=3 D=4 E=5

Catalyst est recodée Catalyst2 : a=1 b=2 c=3 d=4 e=5

justification : obligation d'avoir des valeurs entières pour l'utilisation du module

*Latin-Square et Gréco-Latin de Statistica*

**Data** : 2 interprétations voir 17b) carré latin répliqué sur Time

Voir 17c) carré Gréco-Latin

### Modèles et analyses

**17a)** Modèle M1 :  $Y = \text{Yield}$   $X = \text{Catalyst}$  autres X: rôle de réplication

protocole = plan complètement aléatoire (CRD)

**17b)** Modèle M2 :  $Y = \text{Yield}$   $X = \text{Catalyst}$  facteurs de blocage = Batch, Acid

protocole = plan carré Latin facteur de réplication = Time (ou Time2)

**17c)** Modèle M3 :  $Y = \text{Yield}$   $X = \text{Catalyst}$ , facteurs de blocage = Batch, Acid, Time

protocole = plan Gréco-Latin

**17d)** Comparer les 3 modèles et proposer une conclusion.

### Aide No 17

**17a)** Dans toute analyse statistique en présence de facteurs (variables) qui ne sont pas tenu en compte dans un modèle proposé pour faire l'analyse, a pour conséquence que ces variables jouent un rôle répétition puisque la réponse se répète lorsque ces variables changent et que les variables du modèle sont constantes. En conclusion, les variables qui ne seraient pas significatives dans l'explication (modélisation) de la réponse jouent un autre rôle tout aussi important : elles permettent la réplication des données, ce qui est un principe extrêmement utile pour la validation du modèle en capturant l'erreur expérimentale.

**17b)** la terminologie de facteur (variable) bloc attire l'attention sur la division des variables. Il y a les variables (facteurs) qui présentent l'intérêt premier de la collecte des données. Ce sont les facteurs principaux dont on s'interroge sur leur influence potentielle sur une variable de réponse. Ces facteurs sont à la base du plan expérimental. Mais dans la collecte des données (plan expérimental), il y a toujours d'autres variables secondaires dont il faut tenir compte dans la modélisation. Ce sont les facteurs secondaires dont on doit tenir compte dans le modèle sinon ils viennent gonfler l'erreur expérimentale et fausser les conclusions à propos des facteurs principaux. Quand on les contrôle on leur donne le nom de BLOC. Quand on les mesure on leur donne le nom de COVARIABLE. Peu importe le nom. La caractéristique importante est l'identification du rôle : Principal ou secondaire dans le modèle d'analyse. En particulier, dans le plan fractionnaire carré latin, le nom de **facteur secondaire** peut être remplacé par le terme de **facteur ligne** et par le terme **facteur colonne**.

## Mth8301-exer-18 : Étude sur des médicaments contre la douleur

Données = **Pain.sta** disponible dans les filières  
MTH8301-Data-devoirs.stw ou MTH8301-Data-devoirs.xlsx

### Problématique

Expérience pour comparer l'efficacité de 2 médicaments (Kerlosin, Laposec) contre la douleur.

drug : Kerlosin, Laposec, Placebo (pseudo médicament) est utilisé comme un contrôle.

21 patients (p1, p2, ..., p21) reçoivent l'un des 3 médicaments.

La douleur  $Y_{\text{Pain}}$  est mesurée  $t$  heures après son absorption  $t = 1, 2, 3, 4, 5, 6$

$Y_{\text{Pain}}$  mesure la douleur sur une échelle de 0 (nulle) à 100 (insupportable).

La pression artérielle du sujet est mesurée et on considère que c'est une covariable pouvant avoir une incidence possible sur la mesure de la douleur.

Le fichier contient 126 observations :

21 patients x 6 temps = 7 patients (par médicament) x 3 médicaments x 6 temps

**18a)** Proposer un modèle (équation) pour analyser les données.

Préciser la nature et le rôle des variables.

Ne pas effectuer l'analyse statistique.

**18b)** Mettre en œuvre l'analyse du modèle en **18a)**

Présenter les principaux éléments de votre analyse.

Ne pas faire l'analyse des résidus.

Présenter une conclusion générale sur l'efficacité des médicaments.

Comparer les 2 médicaments Kerlosin et le Laposec.

**18c)** Autre analyse

Régresser  $Y_{\text{Pain}}$  sur la variable pression artérielle ( $X$ ) :  $Y_{\text{Pain}} = \beta_0 + \beta_1 X + \varepsilon$  (\*)

Posons  $Y_{\text{PainPred}}$  : la variable de prédiction du modèle (\*).

Considérer la nouvelle variable  $Y_{\text{PainRes}} = Y_{\text{Pain}} - Y_{\text{PainPred}}$

Refaire l'analyse en vous basant sur  $Y_{\text{PainRes}}$ .

Comparer les résultats de cette analyse avec celle en **18b)**.

### Aide No 18

La structure des données fait appel au concept de mesures répétées (ou longitudinales) car les mêmes patients (sujets, unités expérimentales) sont mesurés à 6 fois dans le temps (heure) ce qui introduit un degré de dépendance dans la mesure de la douleur  $Y$ . Cette dépendance dans les valeurs de la réponse doit être tenu en compte dans l'analyse. Il s'agit de la méthode connue sous le nom de mesures répétées.

Pour contrer la dépendance de la réponse  $Y_{\text{Pain}}$  dans le temps, on considèrera l'analyse de la réponse à 2 périodes seulement :  $Y_{\text{Pain-1}}$  (après 1 heure) et  $Y_{\text{Pain-6}}$  (après 6 heures). Le facteur principal de l'étude est le médicament (Drug) avec 3 modalités (Kerlosin, Laposec, Placebo). Les 21 patients sont divisés en 3 groupes de 7 et dans chaque groupe les patients reçoivent un des 3 médicaments. On est aussi en présence d'une covariable (facteur secondaire) qui est la Pression ( $X$ ) du patient.



## Questions

---

**No18d)** Pour chacune des 2 réponses  $Y_{\text{Pain-1}}$  et  $Y_{\text{Pain-6}}$ , proposer un modèle approprié afin d'analyser si le facteur médicament a une influence significative pour réduire la douleur.

**No18e)** Effectuer une analyse pour valider l'hypothèse qui a été utilisé dans le modèle utilisé dans **No18d)**

**No18f)** Pour la variable  $Y_{\text{Pain-6}}$  seulement, effectuer une régression sur la variable  $X$  (Pression) sur l'ensemble des données en ne tenant pas en compte le facteur Drug. Calculer les valeurs prédites du modèle développé, disons  $Y_{\text{Pain-6-pred}}$  et calculer une nouvelle variable de réponse  $Y_{\text{Pain-6-res}}$  définit par :

$$Y_{\text{Pain-6-res}} = Y_{\text{Pain-6}} - Y_{\text{Pain-6-pred}}$$

Cette nouvelle réponse permet d'enlever l'effet de la covariable  $X$  sur l'ensemble des patients et de considérer tous les patients comme ayant la même pression.

Effectuer une analyse appropriée avec la variable de réponse  $Y_{\text{Pain-6-res}}$  pour décider si le facteur Drug a un effet significatif sur cette nouvelle réponse.

Comparer le résultat de cette analyse avec celle du **No18d)**

## **Mth8301-exer-No 19 : Food Technology**

---

Données = **Food.sta** disponible dans les filières  
MTH8301-Data-devoirs.stw ou MTH8301-Data-devoirs.xlsx

### Référence

A. M. Swanson, J. J. Geissler, P. J. Magnimo, J.A. Swanson, Food Technology vol 21, 1967 p. 99  
The data is extracted from a larger experiment on a sterilized concentrated baby formula.

### Problématique (Purpose)

Select conditions which would provide a formula with acceptable stability for over a year at room temperature, and which would withstand terminal sterilization once the concentrate had been diluted with water.

Design: fractional factorial of 27 trials with 4 factors at 3 values and 1 factor at 5 values  
27 trials =  $2^{5-1}$  points (FF point) + 10 axial points + 1 center point = 16 + 10 + 1 = 27 points  
Response variable Y measured after t months of storage t = 0, 3, 6, 8

Time is a factor.

RunOrder : ordre d'exécution des essais

Response Y : Brookfield viscosity after t = 0, 3, 6, 9 (months)  
measurement unit is in centipoise.

Y0 Y3 Y6 Y9 values of Y at month 0 3 6 9

<u>Facteurs</u>	<u>nom</u>	<u>valeur1</u>	<u>valeur2</u>	<u>valeur3</u>		
A: preheating (deg F) of the milk solids for 25 min	A_preh	175	190	205		
B: addition (%) of sodium polyphosphate	B_sodPOLY	0	0,07	0,14		
C: addition (%) of sodium alginate	C_sodALG	0	0,15	0,30		
D: addition (%) of lecithin	D_lec	0	0,75	1,50		
E: addition (%) of carrageenan	E_carr	0,032	0,135	0,238	0,341	0,444

<u>Variables de codage</u>	<u>valeur1</u>	<u>valeur2</u>	<u>valeur3</u>
XA = (Apreh - 190) / 15	-1	0	1
XB = (B_sodPOLY - 0,07) / 0,07	-1	0	1
XC = (C_sodALG - 0,15) / 0,15	-1	0	1
XD = (D_lec - 0,75) / 0,75	-1	0	1

### QUESTIONS

---

**19a)** Il y a 27 traitements (colonne essai) définis par les combinaisons des 5 facteurs.

Faites un graphique donnant l'allure de Y en fonction du temps (mois) pour les traitements 1 à 9.  
Faire un deuxième graphique pour les traitements 10 à 18 et finalement un troisième graphique pour les traitements 19 à 27.

Donner, si possible, une conclusion générale de la viscosité en fonction du temps qui est indépendante des facteurs.

**19b)** Effectuez une transformation de Box-Cox de Y. Cette nouvelle variable sera notée Ybc. Comparer cette transformation avec la transformation logarithmique logY. Faites un choix entre Ybc et logY et justifier.

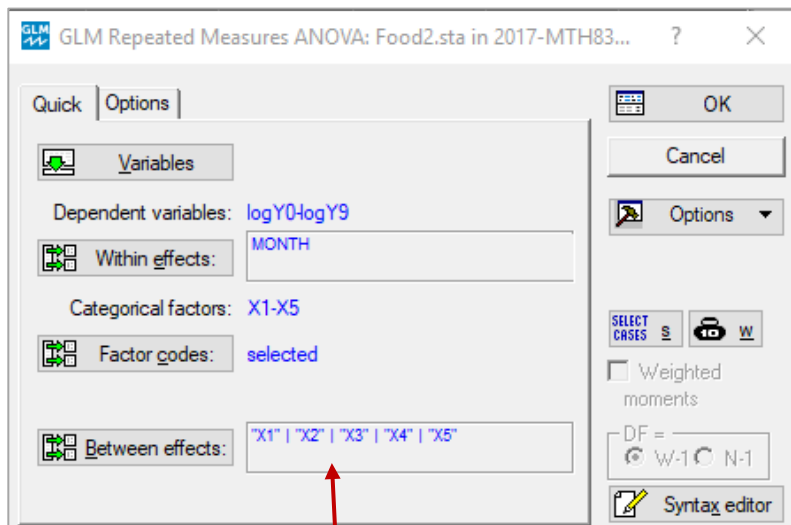
**19c)** Réorganisez les données afin de faire une analyse de logY en *mesures répétées* avec le temps comme facteur de répétition. Présenter un tableau d'analyse de la variance global impliquant les 5 facteurs X1 X2 X3 X4 X5 et le facteur temps. Conclure.

**19d)** À chaque période  $t = 0, 3, 6, 9$ , ajuster un modèle complet du second degré pour la réponse  $Y_t$  en fonction des 5 facteurs. Présenter les coefficients des 4 modèles dans un tableau synthèse. Présenter un tableau d'analyse de la variance pour chaque modèle. Tester si le modèle du premier ordre a un manque d'ajustement.

**19e)** Examiner les résidus selon l'ordre d'exécution des essais, cette information étant disponible dans le fichier. Conclure.

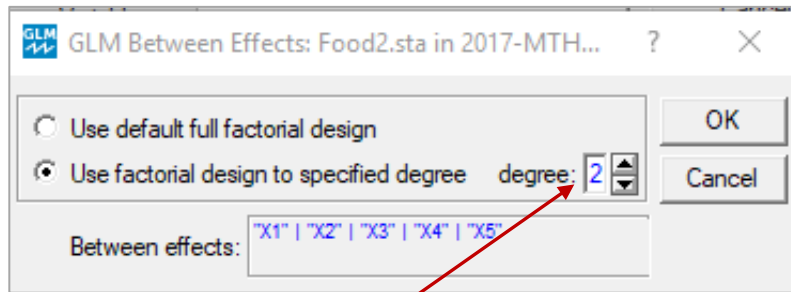
**19f)** Pour être acceptable, la viscosité  $Y_t$  au temps  $t$  ( $t = 0, 3, 6, 9$ ) doit être comprise entre 800 et 1000. Déterminer des conditions pour atteindre cet objectif. Faites des courbes de niveau de la réponse.

### Aide pour la spécification du modèle en No 19c)

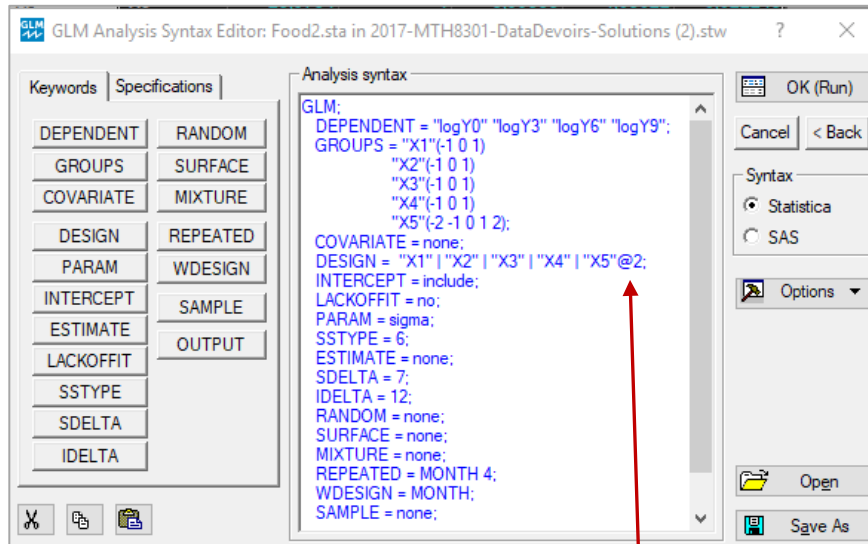


Le modèle (between effects) a trop d'effets : 5 principaux, 20 interactions ordre 1 (produit de 2 facteurs), 20 interactions ordre 2, ..., 1 interaction quintuple ( $X1 \cdot X2 \cdot X3 \cdot X4 \cdot X5$ ) : 31 effets.

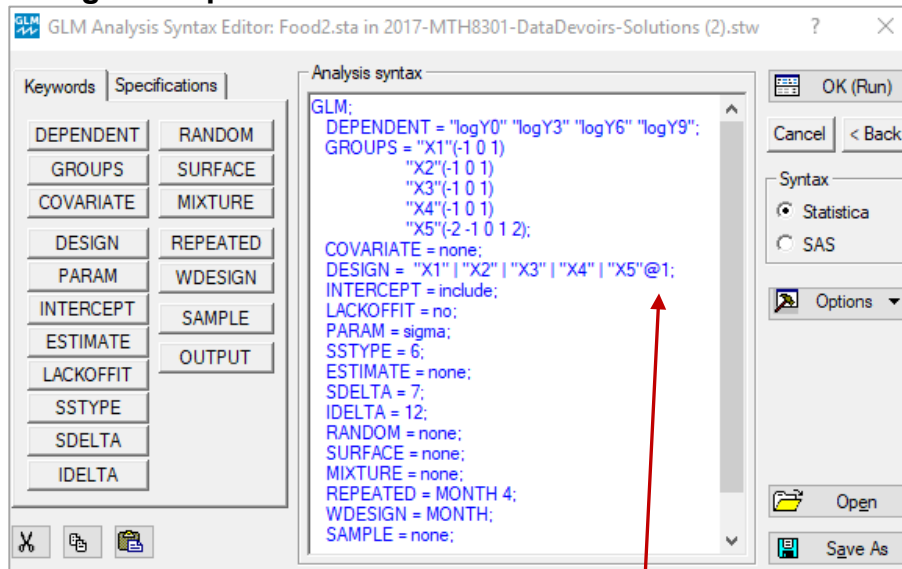
Il n'y pas assez de données pour estimer tous ces 31 effets dont la plupart sont inutiles. Il faut donc proposer est un modèle plus simple. Cliquer ici pour choisir un design au degré 2.



Mais le degré 2 est trop élevé; un modèle avec des effets principaux et interactions doubles est suffisant.



changez le 2 pour 1 donnera le modèle recherché.



Le modèle contiendra seulement des effets principaux des 5 X et le facteur MONTH. Il y aura aussi des interactions de degré 1 des X (= interactions doubles X1\*X2, X1\*X3, etc.) ainsi que des interactions de MONTH avec les 5 X.

## Mth8301-exer-No 20 : Étude de design en génie mécanique

Données = **Hatchback.sta** disponible dans les filières  
MTH8301-Data-devoirs.stw ou MTH8301-Data-devoirs.xlsx

### Problématique

Dans une étude pour cibler la force nécessaire de fermeture du hayon arrière (« hatchback ») dans le design d'une nouvelle voiture, les trois facteurs suivants furent choisis pour les tests :

P = position F = rigidité L = longueur

On décida de faire varier chaque facteur avec 2 modalités : - (-1) et + (+1)

Les valeurs réelles de ces modalités ne sont précisées.

Les voitures (unités expérimentales) présentaient beaucoup de variabilité, il fut décidé de répéter les tests sur chaque voiture. Dans un premier temps, le plan complet de 8 essais avec les 3 facteurs P F L fut retenus. On jugea que 4 tests sur chaque voiture était souhaitable.

Le plan aurait alors demandé 32 essais. Cela demandait trop de temps et était trop couteux. On décida de répéter le plan initial de 8 essais ce qui permettait de répartir les 16 essais sur 4 voitures avec 4 tests sur chaque voiture. Ces conditions de tests sont semblables pour 2 voitures et elles sont différentes pour les 2 autres voitures.

Le plan peut alors être vu comme un plan fractionnaire de 16 essais avec 3 facteurs fixés P F L et un facteur Voiture variant à 4 modalités v1 v2 v3 v4. On peut considérer la colonne 5 Voiture(bloc) selon 2 modèles.

### Modèle 1 : M1 : modèle à effets fixés

Voiture est un facteur FIXÉ de type BLOC avec 4 modalités fixées notées v1 v2 v3 v4.

Il influence additivement la moyenne de Y et il n'a pas d'effet d'interaction avec les 3 autres facteurs P F L.

$$M1 : Y\_Force = \text{général} + \text{Voiture}(\text{fixe}) + P + F + L + PF + PL + FL$$

### Modèle 2 : M2 modèle à effets mixtes

Voiture est un facteur ALÉATOIRE avec 4 voitures v1, v2, v3, v4 représentant l'ensemble de toutes les voitures. Il influence la variance de Y et il n'a pas d'effet d'interaction avec les 3 autres facteurs fixés P F L.

$$M2 : Y\_Force = \text{général} + \text{Voiture}(\text{aléatoire}) + P + F + L + PF + PL + FL$$

### Questions

**20a)** Analyser le modèle M1 avec le module standard des plans à 2 modalités.

**20b)** Analyser le modèle M2 avec le module General Regression Model (GRM).

**20c)** Analyser le modèle M2 avec le module VEPAC.

**20d)** Comparer les résultats des analyses **20a) 20b) 20c)**

## Mth8301-exer-No 21 : composition d'un produit d'éclairage

Données = **Candela.sta** disponible dans les filières  
MTH8301-Data-devoirs.stw ou MTH8301-Data-devoirs.xlsx

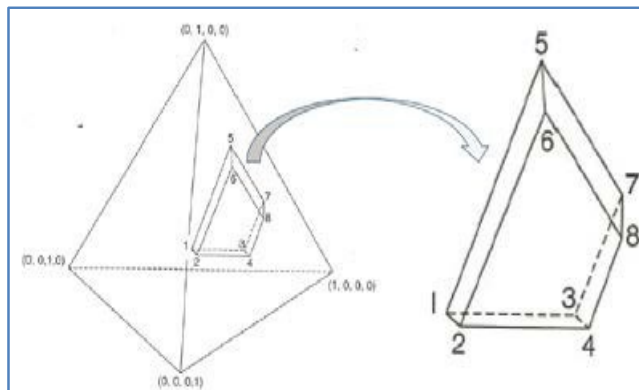
### Problématique

Les constituants chimiques d'un certain produit pour l'éclairage sont :  
magnésium (x1), sodium (x2), nitrate de sodium (x3) et le liant (« binder ») (x4).  
Des études antérieures indiquent que les contraintes de poids suivantes sont appropriées :

$$0,40 \leq x_1 \leq 0,60 \quad 0,10 \leq x_2 \leq 0,50 \quad 0,10 \leq x_3 \leq 0,50 \quad 0,03 \leq x_4 \leq 0,08 \quad (*)$$

La réponse du système est Y : mesure d'illumination (unit = candela).

L'objectif était de trouver un espace expérimental pour réaliser des expériences dont l'analyse des résultats permettrait de déterminer les conditions pour obtenir une réponse maximale. Le plan expérimental choisi, dont la justification n'est pas présentée ici, fut un plan de mélange de 15 essais respectant les contraintes (\*). Le tableau suivant donne la liste des expériences réalisées selon le plan expérimental choisi ainsi que la réponse obtenue Y\_cand. L'espace expérimental est représenté dans le graphique qui suit.



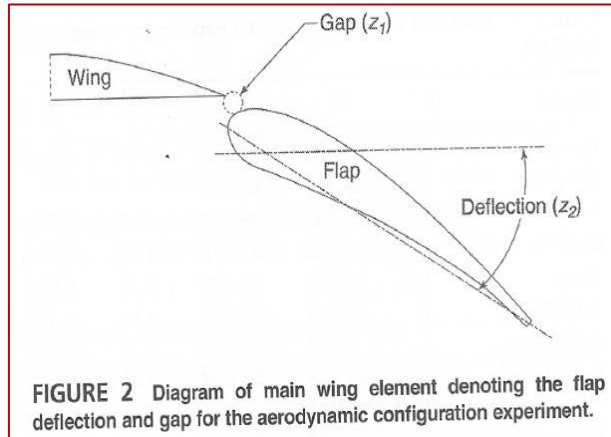
### Questions

- 21a)** Ajustez un modèle quadratique. Commenter les résultats.
- 21b)** Ajustez un modèle du premier ordre. Commenter les résultats.
- 21c)** Présentez un tableau synthèse comparatif des modèles **21a)** et **21b.)**  
avec les coefficients des variables recodées, p-value, R2\_sqr R2\_adj  
Faites un choix entre les 2 modèles et maximiser la réponse.
- 21d)** Commentez l'analyse des résidus du modèle retenu.
- 21e)** Proposez et justifiez 5 essais additionnels à faire.

## Mth8301-exer-No 22 Wind Tunnel Experiment

Données = **AircraftWing.sta** disponible dans les filières  
MTH8301-Data-devoirs.stw ou MTH8301-Data-devoirs.xlsx

Source : Quality Engineering 2007 / Aerodynamic Configuration Aircraft Wing Design



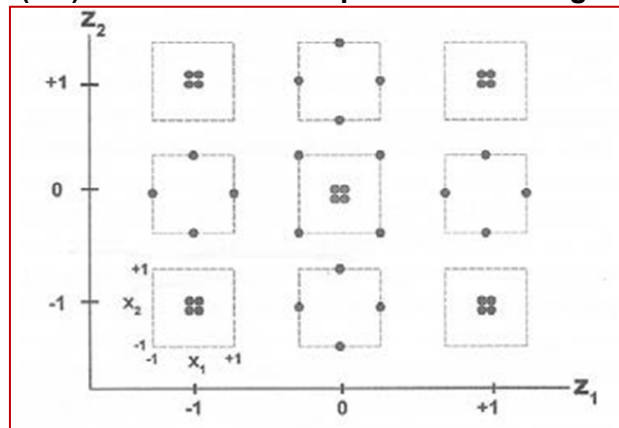
**FDAC** : Facteurs Difficiles A Changer :  $z1\_gap$   $z2\_deflection$

**FFAC** : Facteurs Faciles A Changer :  $x1\_angle$   $x2\_reynold$

Les 4 facteurs  $z1$   $z2$   $x1$   $x2$  sont des facteurs fixes et varient à 3 modalités : -1 0 1

**Réponse** :  $Y\_Lift$

**Plan** : Box-Behnken (BB) de 48 essais tel qu'illustré dans figure



**Whole Plot (WP)** : les unités WP sont identifiées par  $wp1$   $wp2$   $wp3$  ....  $wp12$  et définies par l'intermédiaire des valeurs de  $z1$   $z2$

**WP** constitue un facteur aléatoire

$x1$  et  $x2$  sont des facteurs emboîtés dans WP défini par  $z1$ ,  $z2$

$z1$  et  $z2$  sont des facteurs croisés

il y a 4 essais sur  $x1$  et  $x2$  dans chaque WP  $wp1$   $wp2$  ,....  $wp10$  sont tous distincts

$wp10$   $wp11$   $wp12$  sont identiques et constituent des répétitions.

### Il y a 2 termes d'erreur (sigma)

sigmaWP : l'écart-type des WP (WholePlot) (parcelles)

sigmaSP : l'écart-type des SP (SplitPlot) (parcelles divisées)

Cette quantité correspond à  $\sigma$  (erreur) dans les modèles statistiques sans des effets aléatoires (effets non contrôlés)

### Questions

#### **22a) Analyse des données comme si l'assignation fut en mode complètement aléatoire**

Effectuez l'ajustement d'un modèle polynomial du second degré de Y\_Lift comme si l'assignation des traitements définis par (z1 z2 x1 x2) aurait été faite selon le **mode complètement aléatoire (CRD)**.

On ne tiendra pas en compte de l'assignation réelle qui fut en mode SplitPlot.

Commentez les résultats de l'analyse. Quelles sont les conditions qui maximise Y\_Lift?

**22b)** Refaire l'analyse avec le modèle basé sur les 4 termes linéaires et les 6 termes d'interaction.

#### **22c) Analyse en mode SplitPlot avec VEPAC**

Développer un modèle de Y\_Lift basé sur les 4 effets principaux (z1 z2 x1 x2) et les 6 effets d'interaction  $x1*x2$ ,  $x1*z1$ ,  $x1*z2$ ,  $x2*z1$ ,  $x2*z2$ ,  $z1*z2$ .

Utiliser le module VEPAC avec le modèle suivant :

$$Y = \text{gen} + \text{ID\_WP} + x1 + x2 + z1 + z2 + x1*x2 + x1*z1 + x1*z2 + x2*z1 + x2*z2 + z1*z2 + \text{erreur}$$

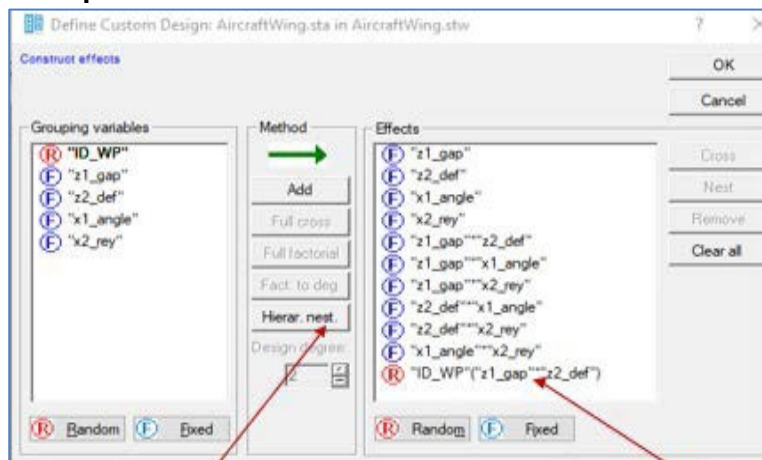
z1, z2, x1, x2 sont des facteurs fixes ID\_WP est un facteur aléatoire.

### Remarques

Il n'y a pas de terme au carré dans le modèle car on considère tous les facteurs comme des variables catégoriques. choisir l'option **type III** pour Sum of Square

Dans VEPAC, on évalue les composants de la variances (ici 2 variances) et on teste l'influence des facteurs fixes sur la réponse. les variables x1 et x2 sont emboîtées dans les variables z1 z2 par l'intermédiaire de ID\_WP

La capture d'écran suivante illustre l'emboitement avec la commande *Hierar.nest*.





**22d)** Présenter plusieurs graphiques de l'analyse **22c)** illustrant la réponse en fonction des effets d'interaction.

**22e)** Comparer les résultats du modèle **22b)** avec ceux du modèle **22a)** concernant les tests de signification des effets. Commentez.

**22f)** Calcul direct de la variance WP et de la variance SP

On peut estimer directement  $\sigma_{WP}^2$  et  $\sigma_{SP}^2$  directement sans tenir en compte  $z_1, x_1, x_2$

C'est une approche de type *erreur pure* qui ne fait pas intervenir la structure SplitPlot du design.

### Étapes

- 1) On calcule les variances  $S_g^2$  de chacun des 7 WP avec des répétitions : wp1, wp2, wp3, wp4, wp10, wp11, wp12.
- 2) L'estimation directe de  $\sigma_{SP}^2$  s'obtient avec la moyenne des 7 variances  $S_g^2$  calculées en 1). Cette estimation est notée par  $s_{SP}^2$ .
- 3) L'estimation directe de  $\sigma_{WP}^2$  s'obtient avec les 3 WP répétés (wp10, wp11, wp12). On calcule les moyennes (notées  $\bar{Y}$ ) basées sur les 4 observations de chacun des 3 WP. On calcule la variance des 3 valeurs  $\bar{Y}$ . Cette variance est notée  $S_{UP}^2$ .  
Estimation de  $\sigma_{WP}^2$  :  $s_{WP}^2 = S_{UP}^2 - (1/4) s_{SP}^2$

Remarque le fichier AircraftWing.sta présente des statistiques descriptives sur les WP

Voir les colonnes 14 à 25 du fichier.

## **Mth8301-exer-No 23 : Expérience multifactorielle**

Données = **Multifacteurs.sta** disponible dans les filières  
MTH8301-Data-devoirs.stw ou MTH8301-Data-devoirs.xlsx

### **Problématique**

Les données proviennent d'une étude pour **évaluer un programme de formation** ainsi qu'une évaluation de candidats pour occuper des postes techniques ainsi que des postes de gestion dans une entreprise de production de produits alimentaires. Les candidats doivent résoudre trois problèmes et on note le nombre de bonnes réponses. Voici une copie des 5 premières lignes du fichier qui en contient 144.

1	2	3	4	5	6	7	8	9	10	11
ID	sujet	genre	groupe	heure	payé	traitement	probleme	rep	stress	Y_br
1	s1	H	E	avant	non	t1	pr1	1	1	12
2	s2	H	E	avant	oui	t2	pr1	1	0	7
3	s3	H	E	après1	non	t3	pr1	1	7	8
4	s4	H	E	après1	oui	t4	pr1	1	1	15
5	s5	H	E	après2	non	t5	pr1	1	7	19

Colonne	Valeurs
ID	1, 2, 3, ..., 144
Sujet	s1, s2, ....., s48
Genre	H, F
Goupe	E=expérimental, C=contrôle
Heure	avant, après1, après2
Payé	oui, non
Problème	pr1, pr2, pr3
Traitement	croisement de genre X groupe X temps X payé : t1 t2 t3 ... t 24
Répétition	(1, 2) du test avec 2 groupes distincts de 24 sujets
Stress	auto mesure de stress par le candidat
Réponse	Y_br : nombre bonnes réponses

- No23a)** Décrire le rôle statistique ou autre de chacune des variables du fichier.  
Rôle statistique : terminologie que vous appris dans le cours MTH8301.  
Soyez le plus spécifique possible.  
Suggestion : présenter le résultat sous forme de tableau.
- No23b)** Réorganiser le fichier pour faire une analyse en mesures répétées.  
Décrire sommairement comment vous avez fait la réorganisation.  
Présenter une copie des 5 premières lignes du nouveau fichier.  
Évaluer le ou les coefficient (s) de corrélation de la réponse.
- No23c)** Faisant abstraction de la structure en mesures répétées, effectuer une analyse statistique de la réponse avec un modèle linéaire comprenant des effets jusqu'à l'ordre 2. Interpréter le résultat et illustrer avec des graphiques pertinents.  
Ne pas faire l'analyse des résidus.
- No23d)** Effectuer une deuxième analyse en tenant en compte la structure en mesures répétées.  
Interpréter le résultat et illustrer avec des graphiques. Ne pas faire l'analyse des résidus.
- No23e)** Comparer les résultats de **No23c)** et **No23d)**.

## **Mth8301-exer-No 24 : Travail de recherche d'une expérience scientifique**

Effectuez une recherche dans une revue scientifique (disponible sur l'internet) afin de trouver un article qui traite d'une expérience réalisée et d'une analyse statistique avec une orientation en ingénierie. L'article doit obligatoirement contenir :

- une analyse statistique
- l'interprétation de l'analyse
- le fichier des données de l'expérience

### **Remarques sur l'article et les données**

Je suggère de contacter les auteurs et de préciser le but de votre démarche : travail de recherche dans un cours de planification d'expériences avec un cas d'application en ingénierie.

En général, ils seront heureux de vous aider en vous communiquant de l'information additionnelle comme l'accès aux données qui ne sont pas toujours accessibles dans l'article. La pratique récemment observée est de les mettre accessibles avec un lien.

Le but de ce travail est d'actualiser la formation reçue dans le cours MTH8301. Un autre but est de lire des articles scientifiques dans des revues avec comité de lecture. Le plus tôt est le mieux dans la formation en ingénierie. Je crois qu'il est souhaitable de commencer cette pratique le plus tôt possible au premier cycle. Il ne faut pas croire que les cours et les laboratoires sont seuls suffisants pour devenir un ingénieur bien formé.

### **Questions**

**No24a)** Présenter sommairement la problématique étudiée ;

**No24b)** Dresser la liste des facteurs X avec leur espace de variation ;

**No24c)** Définir la (les) variable de réponse Y

**No24d)** Présenter sommairement la méthode de mesure de la réponse Y ;

**No24e)** Présenter le plan expérimental employé par les auteurs ;

**No24f)** Présenter le plan employé en format Statistica et commenter le plan employé ;

**No24g)** Refaire l'analyse des données en utilisant le logiciel *STATISTICA* ;  
et comparer le résultat de votre analyse statistique avec l'analyse statistique des auteurs ;

Écrire un rapport sommaire des questions **No24a)** à **No24g)**.

Identifier clairement la référence de l'article ou le site Internet.

Inclure un fichier workbook nommé « FFFFFFFF.stw » de *STATISTICA*

FFFFFFF = votre nom de famille

Le fichier contiendra les données de(s) auteurs (si disponible)  
ainsi que les procédures d'analyse employées à l'étape **No24g)**.

### **Si cela est possible**

suggérer un plan alternatif plus avantageux qui aurait pu être employé selon des critères que vous préciserez.

### Mth8301-exer-No 25 : conception d'un plan fractionnaire

Un appareil est employé pour enlever les particules de débris sur la surface d'un semi-conducteur. Un ingénieur veut déterminer les conditions optimales d'opération afin d'obtenir la plus grande efficacité. Il décide de conduire une série d'essais planifiés et de baser ses recommandations sur les résultats de l'expérience. Comme les essais sont coûteux, il a décidé de ne pas faire un plan factoriel complet. Il vous demande votre avis sur des plans fractionnaires possibles. En vous fournissant la liste des 8 facteurs et leur domaine de variation.

facteur	unité	domaine
A : vitesse au pré-rinçage	Rpm	500 – 1500
B : durée au pré-rinçage	Sec	1 - 25
C : vitesse au rinçage	Rpm	500 – 1500
D : durée au rinçage	Sec	10 – 50
E : vitesse au post-rinçage	Rpm	500 – 1500
F : durée au post-rinçage	Sec	10 – 25
G : vitesse au séchage	Rpm	4000 – 5000
H : durée au séchage	Sec	25 - 50

#### Questions

- 25a)** Proposer un plan fractionnaire avec un nombre minimal d'essais pour obtenir une résolution IV. Expliciter les relations de définition employées.
- 25b)** Expliciter un tableau de tous les contrastes estimables par le plan en retenant que les effets principaux et interactions doubles.
- 25c)** Simplifier le tableau de signification de tous les contrastes si on admet que les seules interactions possibles sont : AB, CD, EF, GH.
- 25d)** Quel serait le niveau de résolution réel compte tenu de cette dernière information ?

### Mth8301-exer-No 26 : conception d'un plan avec contraintes

On veut déterminer un plan expérimental pour 3 facteurs A, B, C afin d'ajuster le modèle

$$Y = \beta_0 + \beta_1 X_A + \beta_2 X_B + \beta_3 X_C + \beta_{11} X_A^2 + \beta_{22} X_B^2 + \beta_{33} X_C^2 + \beta_{12} X_A X_B + \beta_{13} X_A X_C + \beta_{23} X_B X_C + \varepsilon \quad \varepsilon \sim N(0, \sigma^2)$$

Les facteurs doivent satisfaire les contraintes suivantes :

$$0 \leq X_A \leq 10 \quad 0 \leq X_B \leq 10 \quad 0 \leq X_C \leq 10 \quad X_A - 2 X_B \geq 0 \quad X_A + X_B - X_C \geq 0$$

#### Questions

- 26a)** Utiliser le module de *STATISTICA* “*Designs for constrained surfaces and mixtures Design Points (Vertices & Centroids) for Constrained Surfaces and Mixtures*” pour déterminer un plan d'essais incluant seulement des points sommets et des points en 1 dimension « centroids » en 2 dimensions (« vertices and N-dimensional centroids N : 1 through 2 ») Présentez la liste des essais dans l'ordre standard.

- 26b)** Le plan est-il orthogonal ? Justifiez votre réponse.

**Mth8301-exer-No 27 :**