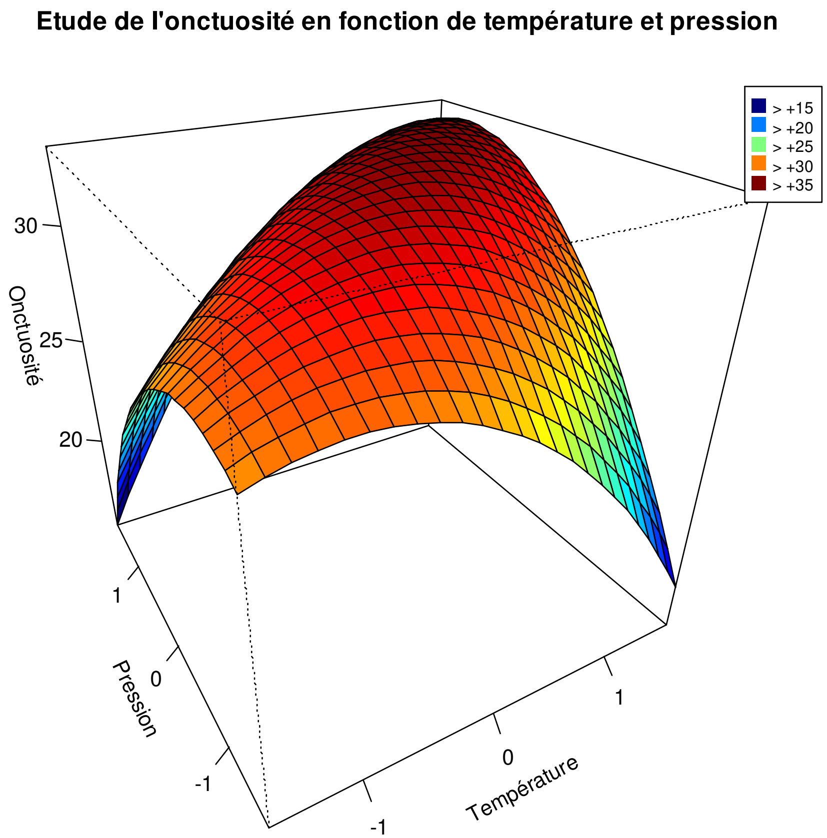
Planification expérimentale

François Husson
UP mathématiques appliquées
Agrocampus Rennes

Quelles expériences réaliser ?



Onctuosite~Temp+Pression+Temp:Pression+I(Temp^2)+I(Pression^2)

Rappels de régression et analyse de variance

$$Y = X\beta + \varepsilon$$

$$\hat{\beta} = (X'X)^{-1} X'Y$$

$$Var(\hat{\beta}) = (X'X)^{-1}\sigma^2$$

Cours sur la planification expérimentale

Les plans fractionnaires

Sir **Ronald Aylmer Fisher**
(1890 – 1962)

Exemple d'utilisation des plans d'expériences

Exemple R&D : modifier la texture de galettes de sarrasin

Objectif : réduire la proportion importante de galettes qui se déchirent lorsqu'on les déplie

Plusieurs variables interviennent dans le process :

- Quantité d'eau (45%, 55%)
- Température de la plaque (180 °, 220 °)
- Étalement de la pâte (automatique, à la main)
- Quantité de pâte par galette (55 g, 65 g)
- Farine (bio, non bio)
- Pliage (à chaud, à froid)
- Température de stockage (6 degrés, 15 degrés)

7 variables
à 2 modalités

Exemple d'utilisation des plans d'expériences

- Quelles expériences réaliser pour déterminer les facteurs influents ?
  - 1ère solution : tester toutes les combinaisons possibles

    $2^7 = 128$ expériences (1 expérience = 1 demi-journée)

    ⟹ Impossible de faire autant d'expériences !!!

- On s'autorise 16 expériences, quel choix faire ?
  - 2ème idée : faire varier 1 facteur à la fois

    ⟹ Pb : impossible d'estimer les interactions

  - 3ème idée : faire varier tous les facteurs à la fois

    ⟹ Difficulté : ne pas confondre les effets des facteurs

Peut-on construire des plans ayant de bonnes propriétés avec peu d'expériences ?

Choix des facteurs et des modalités

On veut généralement :

- étudier le maximum de facteurs
- prendre beaucoup de modalités par facteur

  Pb : nombre d'expériences augmente sensiblement

  Facteurs à 2 niveaux : plans simples mais très utiles car beaucoup d'applications

Les plans complets : matrice des essais

p facteurs à 2 niveaux : toutes les combinaisons sont testées : plan 2^p

Pour 2 facteurs à 2 niveaux : plan 2^2

Matrice des essais :

A	B
+1	+1
+1	−1
−1	+1
−1	−1

- le modèle additif :

$$Y_{ij} = \mu + \alpha_i + \beta_j$$

I	A	B
1	+1	+1
1	+1	−1
1	−1	+1
1	−1	−1

Matrice des effets

- le modèle avec interaction :

$$Y_{ij} = \mu + \alpha_i + \beta_j + \alpha\beta_{ij}$$

I	A	B	AB
1	+1	+1	+1
1	+1	−1	−1
1	−1	+1	−1
1	−1	−1	+1

Les plans complets : matrice des effets

- le modèle additif :

$$\mathbf{X}= \begin{bmatrix} 1 & +1 & +1 \\ 1 & +1 & -1 \\ 1 & -1 & +1 \\ 1 & -1 & -1 \end{bmatrix} \qquad (X'X) = \begin{pmatrix} 4 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 4 \end{pmatrix} = n\ I_3$$

- le modèle avec interaction :

$$\mathbf{X}= \begin{bmatrix} 1 & +1 & +1 & +1 \\ 1 & +1 & -1 & -1 \\ 1 & -1 & +1 & -1 \\ 1 & -1 & -1 & +1 \end{bmatrix} \qquad (X'X) = \begin{pmatrix} 4 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 4 \end{pmatrix} = n\ I_4$$

$(X'X) = n\ Id$ (avec n = nb d'expériences) : matrice d'Hadamard

Qu’est ce qu’un bon plan ?

Qu'est ce qu'un bon plan ?

Choisir les essais qui permettent d'avoir une estimation des effets de chaque variable la plus précise possible

Il faut minimiser : $V(\hat{\beta}) = (X'X)^{-1} \sigma^2$

Dépend uniquement du **choix** des expériences

Variabilité résiduelle : dépend des résultats des expériences

Objectif des plans : trouver les expériences telles que $(X'X)^{-1}$ soit « minimale »

Plan à 3 facteurs en 4 essais

Plan complet 2^3, modèle additif

$$(X'X) = n\ I_4 = 8\ I_4 \qquad (X'X)^{-1} = \begin{array}{c} \begin{array}{cccc} \mathbf{I} & \mathbf{A} & \mathbf{B} & \mathbf{C} \end{array} \\ \begin{pmatrix} 0.125 & 0 & 0 & 0 \\ 0 & 0.125 & 0 & 0 \\ 0 & 0 & 0.125 & 0 \\ 0 & 0 & 0 & 0.125 \end{pmatrix} \end{array} \begin{array}{c} \mathbf{I} \\ \mathbf{A} \\ \mathbf{B} \\ \mathbf{C} \end{array}$$

4 essais choisis au hasard

$$(X'X)^{-1} = \begin{array}{c} \begin{array}{cccc} \mathbf{I} & \mathbf{A} & \mathbf{B} & \mathbf{C} \end{array} \\ \begin{bmatrix} 0.50 & 0.00 & -0.25 & 0.25 \\ 0.00 & 0.50 & 0.25 & 0.25 \\ -0.25 & 0.25 & 0.50 & 0.00 \\ 0.25 & 0.25 & 0.00 & 0.50 \end{bmatrix} \end{array}$$

4 essais bien choisis :

$$\begin{array}{c} \begin{array}{cccc} \mathbf{I} & \mathbf{A} & \mathbf{B} & \mathbf{C} \end{array} \\ \begin{bmatrix} 0.25 & 0.00 & 0.00 & 0.00 \\ 0.00 & 0.25 & 0.00 & 0.00 \\ 0.00 & 0.00 & 0.25 & 0.00 \\ 0.00 & 0.00 & 0.00 & 0.25 \end{bmatrix} \end{array}$$

Remarque : $(X'X)^{-1} = \frac{1}{4}\mathrm{Id}$

Variance de l'estimateur de l'effet du facteur A augmente

Il n'y a plus indépendance entre l'estimation du facteur A et celle du facteur C

Attention : Supprimer des essais au hasard déséquilibre tout

Exercice : plan à 4 facteurs en 8 essais

Construire un plan à 4 facteurs à 2 niveaux en 8 essais

Les 4 facteurs à 2 modalités sont :

- épaisseur de la pâte (fine, épaisse)
- type de farine (bio, non bio)
- pliage de la pâte (à froid, à chaud)
- mode de pliage (automatique, manuelle)

La variable à expliquer est le pourcentage de galette qui se déchirent lorsqu'on les déplie

Télécharger le fichier suivant, puis jouer sur les essais pour retrouver la précision maximum sur l’estimation des paramètres du modèle

https://husson.github.io/img/planfra_4facteurs.xlsx

Construction d'un plan 2^{3-1}

3 facteurs à 2 modalités

en 2^{3-1} = 4 essais

A	B	C
1	1	1
1	1	-1
1	-1	1
1	-1	-1
-1	1	1
-1	1	-1
-1	-1	1
-1	-1	-1

Choix de 4 essais

1ère idée : pour chaque facteur, tester les niveaux 1 et -1 un même nb de fois

2ème idée : pour chaque couple de 2 facteurs, prendre autant de combinaisons (1,1), (-1,1), (1,-1) et (-1,-1)

Choix d'essais dans le cas général :

1ère idée : les niveaux de chaque facteur testés un même nb de fois

2ème idée : prendre autant de combinaisons (1,1), (-1,1), (1,-1) et (-1,-1) pour chaque couple de 2 facteurs

3ème idée : prendre autant de combinaisons (1,1,1), (1,1,-1), (1,-1,1), … pour chaque triplet de 3 facteurs

4ème idée : prendre autant de combinaisons (1,1,1,1), (1,1,1,-1), (1,1,-1,1), … pour chaque quadruplet de 4 facteurs

…

Beaucoup trop compliqué de construire un plan de cette façon
Besoin d'un principe de construction simple si beaucoup de facteurs

Constat

- Un plan complet permet d'estimer tous les facteurs et toutes les interactions d'ordre 2, 3, 4, …
- Interactions d'ordre 3 et + sont souvent négligeables
- Exemple : plan 2^5 :

$$\underbrace{1}_{\text{Constante}} + \underbrace{5}_{\text{Effet}} + \underbrace{10}_{\text{Interactions d'ordre 2}} = 16$$ paramètres intéressants à estimer

$\Rightarrow$ Dommage de faire 32 expériences pour n'estimer « que » 16 paramètres

Principe de construction des plans fractionnaires 2^{p-k}

1. Choix d'un plan de base à 2^{p-k} essais

2. Construction de la matrice des effets du modèle saturé avec ce plan de base

3. Choix des confusions : affectation des effets principaux

4. Détermination des confusions résultantes

Retour sur le plan fractionnaire 2^{3-1}

1. Choix d'un plan de base
   à $2^2 = 4$ essais

2. Construction de la matrice
   des effets du modèle saturé
   avec ce plan de base

3. Le facteur C est confondu avec
   l'interaction AB

4. Détermination des confusions
   résultantes : C = AB

			C
I	A	B	AB
1	1	1	1
1	1	-1	-1
1	-1	1	-1
1	-1	-1	1

Confusion d'effet (alias) et générateur d'alias

$$C = AB \implies CC = ABC \implies I = ABC$$

I	**A**	**B**	**AB**
ABC	BC	AC	C
1	1	1	1
1	1	-1	-1
1	-1	1	-1
1	-1	-1	1

C	**x**	**C**		**I**
1	x	1	=	1
(-1)	x	(-1)	=	1
(-1)	x	(-1)	=	1
1	x	1	=	1

$$I = ABC \implies A(I) = A(ABC) \implies A = BC$$

$$\implies B(I) = B(ABC) \implies B = AC$$

Confusion d'effet (plan 2^{3-1})

Générateur d'alias : I = ABC

X

$$
\begin{array}{cccccccc}
\mathbf{I} & \mathbf{A} & \mathbf{B} & \mathbf{C} & \mathbf{AB} & \mathbf{AC} & \mathbf{BC} & \mathbf{ABC}
\end{array}
$$

$$
\begin{pmatrix}
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 \\
1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 \\
1 & -1 & -1 & 1 & 1 & -1 & -1 & 1
\end{pmatrix}
$$

$X'X$

$$
\begin{array}{c|c}
\begin{array}{cccccccc} \mathbf{I} & \mathbf{A} & \mathbf{B} & \mathbf{C} & \mathbf{AB} & \mathbf{AC} & \mathbf{BC} & \mathbf{ABC} \end{array} & \\
\begin{pmatrix}
4 & 0 & 0 & 0 & 0 & 0 & 0 & 4 \\
0 & 4 & 0 & 0 & 0 & 0 & 4 & 0 \\
0 & 0 & 4 & 0 & 0 & 4 & 0 & 0 \\
0 & 0 & 0 & 4 & 4 & 0 & 0 & 0 \\
0 & 0 & 0 & 4 & 4 & 0 & 0 & 0 \\
0 & 0 & 4 & 0 & 0 & 4 & 0 & 0 \\
0 & 4 & 0 & 0 & 0 & 0 & 4 & 0 \\
4 & 0 & 0 & 0 & 0 & 0 & 0 & 4
\end{pmatrix} & \begin{array}{c} I \\ A \\ B \\ C \\ AB \\ AC \\ BC \\ ABC \end{array}
\end{array}
$$

X'X non inversible car confusion entre I et ABC, entre A et BC, …

Mais si on se restreint à l'étude des effets principaux :

$X'X$ **s'écrit simplement et est facilement inversible :** $(X'X)^{-1} = \frac{1}{n}\mathrm{Id}$

Construction d'un plan fractionnaire 2^{4-1}

1. Choix d'un plan de base
   à $2^3 = 8$ essais

2. Construction de la matrice
   des effets du modèle saturé
   avec ce plan de base

3. L'interaction ABC certainement
   négligeable : confondre le facteur
   D avec l'interaction ABC

4. Détermination des confusions
   résultantes : D = ABC

							D
I	**A**	**B**	**C**	AB	AC	BC	ABC
1	**1**	**1**	**1**	1	1	1	1
1	**1**	**1**	**-1**	1	-1	-1	-1
1	**1**	**-1**	**1**	-1	1	-1	-1
1	**1**	**-1**	**-1**	-1	-1	1	1
1	**-1**	**1**	**1**	-1	-1	1	-1
1	**-1**	**1**	**-1**	-1	1	-1	1
1	**-1**	**-1**	**1**	1	-1	-1	1
1	**-1**	**-1**	**-1**	1	1	1	-1

Confusion d'effet (alias) et générateur d'alias

$$D = ABC \implies DD = ABCD \implies I = ABCD$$

I	**A**	**B**	**C**	AB	AC	BC	ABC		
ABCD	BCD	ACD	ABD	CD	BD	AD	**D**	**D**	**I**
1	1	1	1	1	1	1	1	1	1
1	1	1	-1	1	-1	-1	-1	-1	1
1	1	-1	1	-1	1	-1	-1	-1	1
1	1	-1	-1	-1	-1	1	1	1	1
1	-1	1	1	-1	-1	1	-1	-1	1
1	-1	1	-1	-1	1	-1	1	1	1
1	-1	-1	1	1	-1	-1	1	1	1
1	-1	-1	-1	1	1	1	-1	-1	1

$$I = ABCD \implies A(I) = A(ABCD) \implies A = BCD$$

Construction d'un plan fractionnaire 2^{5-2}

1. Choix d'un plan de base à $2^3 = 8$ essais

2. Construction de la matrice des effets du modèle saturé avec ce plan de base

3. Affectation des effets principaux

4. Détermination des confusions résultantes

				D	E		
I	**A**	**B**	**C**	AB	AC	BC	ABC
1	**1**	**1**	**1**	1	1	1	1
1	**1**	**1**	**-1**	1	-1	-1	-1
1	**1**	**-1**	**1**	-1	1	-1	-1
1	**1**	**-1**	**-1**	-1	-1	1	1
1	**-1**	**1**	**1**	-1	-1	1	-1
1	**-1**	**1**	**-1**	-1	1	-1	1
1	**-1**	**-1**	**1**	1	-1	-1	1
1	**-1**	**-1**	**-1**	1	1	1	-1

D = AB E = AC

Confusion d'effet (alias) et générateur d'alias

$$D = AB \quad \Longrightarrow DD = ABD \quad \Longrightarrow I = ABD$$
$$E = AC \quad \Longrightarrow EE = ACE \quad \Longrightarrow I = ACE$$

On a aussi $E = BCD \quad \Longrightarrow EE = BCDE \quad \Longrightarrow I = BCDE$

$$I = ABD = ACE \quad \Longrightarrow II = (ABD)(ACE) \quad \Longrightarrow I = BCDE$$

I	A	B	C	AB	AC	BC	ABC
ABD	BD	AD	ABCD	D	BCD	ACD	CD
ACE	CE	ABCE	AE	BCE	E	ABE	BE
BCDE	ABCDE	CDE	BDE	ACDE	ABDE	DE	ADE

Confusion d'effets : estimation de paquets d'effets ou interactions.

Paquet bleu estimable mais impossible de savoir ce qui est dû à C, à l'interaction ABCD, l'interaction AE, l'interaction BDE

Nombre de facteurs et nombre d'essais

Résolution = longueur du plus petit générateur d'alias

Exemple : plan 2^{4-1} : I = ABCD Résolution IV

plan 2^{5-2} : I = ABD = BCE = BCDE Résolution III

Résolution III : effet principaux confondus avec interactions d'ordre 2 ou plus

Résolution IV : effet principaux confondus avec interactions d'ordre 3 ou plus

Résolution V : effet principaux confondus avec interactions d'ordre 4 ou plus
et interactions d'ordre 2 confondues avec interactions d'ordre 3 ou plus

s	3	4	5	6	7	8	9
Nb d'expériences : 2^s	8	16	32	64	128	256	512
Nb de facteurs en résolution 3 : 2^s-1	7	15	31	63	127	255	511
Nb de facteurs en résolution 4 : 2^{s-1}	4	8	16	32	64	128	256
Nb de facteurs en résolution 5	3	5	6	8	11	17	≥ 23

Construction de plans avec R (package FrF2)

```
library(FrF2)
plan1 <- FrF2(nruns=8, nfactors=4, factor.names=list(temp=c("min","max"),
     press=c("low","normal"),material=c("M1","M2"),state=c("new","aged")))
plan2 <- FrF2(nfactors=5, resolution=5)
summary(plan2)
```

```
Call:
FrF2(nfactors = 5, resolution = 5)

Experimental design of type  FrF2
16  runs

Factor settings (scale ends):
   A  B  C  D  E
1 -1 -1 -1 -1 -1
2  1  1  1  1  1

Design generating information:
$legend
[1] A=A B=B C=C D=D E=E

$generators
[1] E=ABCD


Alias structure:
[[1]]
[1] no aliasing among main effects and 2fis
```

(suite des résultats)

```
The design itself:
    A  B  C  D  E
1   1  1  1 -1 -1
2   1  1 -1  1 -1
3   1 -1  1 -1  1
4   1 -1 -1  1  1
5  -1  1 -1  1  1
6  -1 -1 -1 -1  1
7   1 -1  1  1 -1
8  -1  1  1  1 -1
9  -1 -1  1 -1 -1
10 -1 -1  1  1  1
11  1 -1 -1 -1 -1
12  1  1  1  1  1
13 -1  1 -1 -1 -1
14  1  1 -1 -1  1
15 -1  1  1 -1  1
16 -1 -1 -1  1 -1
class=design, type= FrF2
```

Dépouillement des résultats

Règles :

On considère négligeables :

1. tous les termes d'un paquet lorsque le paquet est négligeable
2. les interactions d'ordre supérieur ou égal à 3
3. les interactions entre 2 effets négligeables
4. les interactions comprenant un effet négligeable
5. toutes les interactions

Contraintes de + en + fortes

Dépouillement des résultats : exemple plan 2^{5-2}

I	**A**	**B**	**C**	AB	AC	BC	ABC
ABD	BD	AD	ABCD	**D**	BCD	ACD	CD
ACE	CE	ABCE	AE	BCE	**E**	ABE	BE
BCDE	ABCDE	CDE	BDE	ACDE	ABDE	DE	ADE
2.25*	-3.15*	0.35	-0.52	2.58*	4.15*	-0.63	0.26

Les paquets **3**, **4**, **7** et **8** sont négligeables

Règle 1 : tous les termes des paquets **3**, **4**, **7** et **8** sont négligeables

Règle 2 : les interactions d'ordre supérieur à 2 sont négligeables

Règle 3 : les interactions entre 2 effets négligeables sont négligeables (aucune)

Règle 4 : les interactions comprenant un effet négligeable (BD, CE, AB, AC)

Règle 5 : toutes les interactions sont négligeables (inutile)

De la résolution 3 à la résolution 4

- Ajout du plan complémentaire au plan de résolution 3
- Exemple : plan 2^{5-2}

Plan initial :

$$\left.\begin{array}{l} D = AB \\ E = AC \end{array}\right\} I = ABD = ACE = BCDE$$

Plan complémentaire :

$$-D = (-A)(-B) \implies D = -AB$$

$$-E = (-A)(-C) \implies E = -AC$$

Plan complet :

$$\left.\begin{array}{l} D = ABS \\ E = ACS \end{array}\right\} I = ABDS = ACES = BCDE$$

S	A	B	C	D	E	
1	1	1	1	1	1	**Plan initial**
1	1	1	-1	1	-1	
1	1	-1	1	-1	1	
1	1	-1	-1	-1	-1	
1	-1	1	1	-1	-1	
1	-1	1	-1	-1	1	
1	-1	-1	1	1	-1	
1	-1	-1	-1	1	1	
-1	-1	-1	-1	-1	-1	**Plan complémentaire**
-1	-1	-1	1	-1	1	
-1	-1	1	-1	1	-1	
-1	-1	1	1	1	1	
-1	1	-1	-1	1	1	
-1	1	-1	1	1	-1	
-1	1	1	-1	-1	1	
-1	1	1	1	-1	-1	

Démarche statistique

1. Définir la problématique
2. Choisir les expériences à réaliser (planification expérimentale)
3. Effectuer les expériences
4. Dépouiller les résultats (analyse de variance)

Retrouver ce cours sur Youtube

- https://www.youtube.com/HussonFrancois

- Dans Google, taper les mots clés :
  Youtube plans d’expériences Husson