MRE m'a tuer (AZ)

Méthodologie de la recherche expérimentale (agro-zootechnique)

Module 1 : Blocs complets à 1 facteur

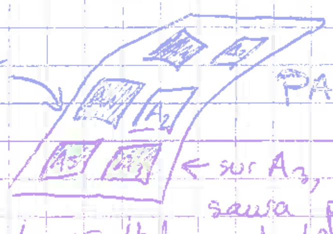
Enregistrement Audio

L'année dernière, on avait 1 facteur étudié à dispositif aléatoire
exemple : facteur A, 3 modalités, 2 répétitions
et on étudiait le CM_E (résiduel) et CM_A,
avec un CM_E le + petit possible

Là on pourra avoir un facteur d'hétérogénéité mais contrôlé (par bloc)

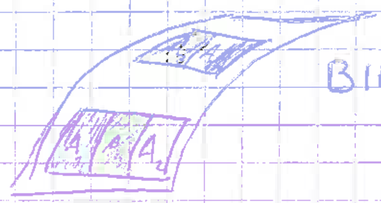
exemple sur une pente physique

parcelle en pente



PAS BIEN

← sur A_3, on ne saura pas si les résultats sont dûs à la pente ou au traitement



BIEN !!

$$\underset{①}{x_{ij}} = \underset{②}{\mu} + \underset{③}{\alpha_i} + \underset{④}{\beta_j} + \underset{⑤}{\varepsilon_{ij}}$$

① réponse pour la modalité i du facteur A dans le bloc j
② niveau moyen
③ effet principal du facteur A de modalité i ($\alpha = 1, 2, 3 \dots a$)
④ effet bloc j ($j = 1 \dots r$)
⑤ effet résiduel pour la mod i du facteur A dans le bloc j

Méthode → calcul de la liste des résidus e_{ij} $a \times r$ résidus

→ analyse des résidus $\mathcal{L}(e_{ij}) = \mathcal{N}(0; \sigma_\varepsilon)$ ≃ 17 min
- histogramme (est-il symétrique ?)
- coeff. de Pearson β_1 et β_2

≃ 21 → indpdce des résidus
- cartographie

23:30 → homoscédasticité
2 variances → F bilatéral test
+ de 2 variances → χ^2 unilatéral test

Quand on a 1 facteur étudié et 1 fact contrôlé, faut 2 tests.
Un chacun quoi.

→ ANOVA • dans 1 fact étu 1 fact contrôlé
↳ 2 paires d'hypothèse (H_0 ; H_1)
↳ 2 tests F unilatéraux

→ ppas + petite amplitude significative

moyennes classées par grandeur et prises par groupes de 5 puis de 4, puis 3, puis 2 ...

$m_1 \quad m_2 \quad m_3 \quad m_4 \quad m_5 \quad m_6$

$(m_5 - m_1) \leq$ ou $>$ à ppas ?

$(m_6 - m_2) \leq$ ou $>$ à ppas ?

$(m_4 - m_1) \leq$ ou $>$ à ppas ? etc.

ppas calculé avec ddle

$$ppas = q_{1-\alpha} \times \sigma(m) \qquad \sigma(m) = \sqrt{\frac{CMe}{nb\ bloc}}$$

n est le nb de valeurs qui a servi à calculer m_i

table Newman Keuls

en zootechnie, y a toujours un pb avec le poids, alors on les classe les animaux par catégories.

Question

4 facteurs : A → 4 modalités
B → 16 modalités

N° loge 1 2 3 4 5

Loge n	1	2	3	4	1
	1	1	1	1	2
	2	2	4	1	2
	2	2	2	3	3

Dans les loges (de 5), les blocs (de 4) sont mélangés

Module 2 : Blocs à 2 facteurs et plus

On a un facteur étudié A (a modalités) $i = 1 \ldots a$
un facteur étudié B (b modalités) $j = 1 \ldots b$
un facteur contrôlé, bloc de r blocs $k = 1 \ldots r$

Nombre de modalités $= a \times b \times r$
Nombre de trait^{t} $= a \times b$ répétés r fois

Ainsi, on devrait avoir $x_{ijk} = \mu + \alpha_i + \beta_j + \gamma_k + \alpha_i\beta_j + \varepsilon_{ijk}$

Pas d'interaction entre bloc et facteurs étudiés.

Par exemple $\left.\begin{array}{l} a = 2 \\ b = 3 \end{array}\right\}$ 6 traitements
$r = 4$

			A_1	B_2		bloc 1	bloc 2
bloc 1	a_1b_1 a_2b_2 a_1b_2 a_2b_3 a_1b_3 a_2b_1	B_1			$\bar{x}_{B_1}$	moy$_1$	moy$_2$
bloc 2		B_2			$\bar{x}_{B_2}$		
bloc 3		B_3			$\bar{x}_{B_3}$		
~~bloc 4~~			$\bar{x}_{A_1}$	$\bar{x}_{A_2}$			

Note : pour les résidus, on a la formule (démontrée [illegible] plus)
formule à connaître

B.B " Je vous demande plus de savoir interpréter que calculer "

1000^e RAPPEL → Vérifier la normalité (Pearson, carto, histo)
→ Indpdce des résidus
→ Homoscédasticité (4 tests → 1 pour A, 1 pour B, 1 pour bloc, 1 pour AB)

Homoscédasticité

2 variances → test F
+2 variances → χ^2 unilat

Pour le facteur A (2 variances) → F $\quad F_{calc} = \dfrac{SCE_{A_1} / \nu_1}{SCE_{A_2} / \nu_2}$

avec $\nu_1 = 11$ (on a 12 ~~nom~~ réponses pour chaque fact)
$\nu_2 = 11$

Pour le facteur B (ici 3 variances) → χ^2

pour $\chi^2_{théo}$, ddl = nb variances − 1 = 2

pour χ^2_{calc}

Sur 24 valeurs [illegible] 3 variances

n_j	ν_j	$1/\nu_j$
8	7	1/7
8	7	1/7
8	7	1/7
24	$\nu = 21$	3/7

$\sum \frac{1}{\nu_j} - \frac{1}{\nu} = \frac{3}{7} - \frac{1}{21}$

bloc → χ^2 (4 variances) $\chi^2_{1-\alpha}(3)$

n_j	ν_j	$1/\nu_j$
6	5	5
6	5	5
6	5	5
6	5	5
24	20	4/5

$$\sum \frac{1}{\nu_j} - \frac{1}{\nu} = \frac{4}{5} - \frac{1}{20}$$

Interaction A×B → χ^2 (6 variances) $\chi^2_{1-\alpha}(5)$

n_j	ν_j	$1/\nu_j$
4	3	1/3
4	3	1/3
4	3	"
4	3	"
4	3	"
4	3	"
24	18	2

$$\sum \frac{1}{\nu_j} - \frac{1}{\nu} = 2 - \frac{1}{18}$$

ANOVA

ddl de ddl :

	SCE	ddl	CM	Fcalc	$F_{1-\alpha}(\nu_1 ; \nu_2)$
A	$SP_A - C$	$a-1$	$\frac{SCE}{ddl}$	4 tests F	
B	$SP_B - C$	$b-1$	"		
A×B	$SP_{A\times B} - C - SCE_A - SCE_B$	$(a-1)(b-1)$	"		
bloc	$SP_{bloc} - C$	$r-1$	"		
e	(par différence)	(par dif)	"		
ToT	$SC - C$	$abr-1$			

PPAS

~~si il n'y a pas~~ d'Interaction AB significative ?

- non → Effet A significatif ?
  - oui → ppas ddle, $n = b\times r$
  - non → rien à faire
- non → Effet B sign ?
  - oui → ppas ddle, $n = a\times r = 8$ *
  - non → rien à faire
- oui → ppas sur les a×b moyennes (sur les a×b traitement), $n = r$

⚠ PPAS $q^k_{1-\alpha}(\nu)\sqrt{\frac{CM_e}{n}}$

* raisonnement : on a utilisé 24 valeurs pour 3 variances $\frac{24}{3} = 8$

Module 3 : Carré latin et gréco-latin

1 facteur étudié et 2 sources d'hétérogénéité

On considère que pour chaque source d'hétérogénéité, on a 4 modalités.
On a donc 4 lignes et 4 colones

Le carré latin est alors comme ça (on considère que le facteur étudié a 4 modalités, donc on a 4 traitements)

	c_1	c_2	c_3	c_4
l_1	~~T_2~~	T_4	~~T_1~~	~~T_3~~
l_2	T_4	~~T_3~~	~~T_2~~	~~T_1~~
l_3	T_1	~~T_2~~	~~T_3~~	T_4
l_4	~~T_3~~	~~T_1~~	T_4	~~T_2~~

Modèle

$$x_{ijk} = \mu + \alpha_i + \beta_j + \gamma_k + e_{ijk}$$

avec $\alpha_i = \bar{x}_{i..} - \mu$ $\quad \beta_j = \bar{x}_{.j.} - \mu$

$\gamma_k = \bar{x}_{..k} - \mu$

$p = 4$ (trait^{nt})
$N = 16$ (unités expé)

$$e_{ijk} = x_{ijk} - \mu - \alpha_i - \beta_j - \gamma_k$$
$$= x_{ijk} - \mu - \bar{x}_{i..} + \mu - \bar{x}_{.j.} + \mu - \bar{x}_{..k} + \mu$$

$$e_{ijk} = x_{ijk} + 2\mu - \bar{x}_{i..} - \bar{x}_{.j.} - \bar{x}_{..k}$$

On aura 3 tests de χ^2. Le ddl pour $\chi^2_{1-\alpha}$ est $p-1$

pour le χ^2_{calc}, on aura $\sum \frac{1}{\nu_j} - \frac{1}{\nu} = p \times \frac{1}{p-1} - \frac{1}{p(p-1)}$
$= 4 \times \frac{1}{3} - \frac{1}{12}$

ANOVA

facteur A	SCE	ddl	CM	F
étudié		$p-1$		$F_{étu}$
Ligne		$p-1$		F_{lig}
colonne		$p-1$		F_{col}
résidu		p^2-3p+2		
total		p^2-1		

c'est bien entendu le facteur étudié qui nous intéresse

ppas

$$ppas = q^{k}_{1-\alpha}(\nu)\sqrt{\frac{CM_e}{n}}$$

NB : si on a 2 moyennes de même valeur, on met bien les 2 dans le test ppas

NB : Le carré latin, c'est donc un carré (côtés même longueur = même nb facteurs)
spécificité → 1 seule fois un chiffre par colonne
ligne (sudoku!)

MRE3

Le carré gréco latin

Quand tu as plusieurs facteurs étudiés, tu fais 2 carrés latin et tu les fusionne ! (POWER !)

Agro alim

Modules 1, 2, 3, ~~4, 7~~, 8 (Bande avancée)

Rappel partie 7 de 2A

On considérait des variables naturelles u : $B = (u'u)^{-1} u'y$
codées x : $B = (X'X)^{-1} X'y$

codée → critère qualitatif donné en chiffre ~~ex~~ :
ex : lait vache → 0
lait chèvre → 1

On faisait aussi l'ANOVA et le test sur les coef.

Cette année, on va construire le dispositif expé optimum $B = (X'X)^{-1} X'y$

Dans l'ANOVA, la SCE_e sera divisée en 2 : → SCE_{ad} → $SCE_{répet}$

Pour valider un modèle prévisionnel et/ou structurel

Pour le test sur coef, le ddl

$Var B = \sigma_\varepsilon^2 (X'X)^{-1}$ ddl devient → CM_e
soit ↳ $CM_{répet}$ ⟸ [illegible]

Les modules 9 et 10 concerneront l'optimisation

Module 1 : Plan de criblage

$$\hat{y} = \beta_0 + \sum_1^k \beta_j X_j = \beta_0 + \beta_1 X_1 + \beta_2 X_2 [\ldots] + \beta_k X_k$$

On a donc k facteurs à cribler
Chaque facteur est à 2 niveaux (1 ou −1)

Mettons : 200°C → 1
240°C → −1
220°C → 0

Si on a 6 facteurs à 2 modalité, théoriquement on pourrait penser qu'il faut 2^6 expériences.
Mais on peut le faire avec 8 expériences seulement !!!

La matrice expé d'un plan de criblage : $\hat{y} = \beta_0 + \sum_1^k \beta_j X_j$

On soumet la 1er ligne de la matrice à l'algorithme de Plackett et Burman

Pour k' coef il faut k' essais. coef → inconnus
essais → équations

exemple (facteurs à 2 mod) et 6 facteurs

$k = 6$
$k' = 7$ } algorithme $N = 8$ (normal)

On prend une ligne qu'on recopie en décalant vers la droite (oui, oui)

<u>Illustration</u>

La 1ère ligne de la matrice est donnée par l'algorithme Placket et Burman

+	+	+	−	+	−	−	
−	+	+	+	−	+	−	(le dernier revient au début)
−	−	+	+	+	−	+	re
+	−	−	+	+	+	−	re
−	+	−	−	+	+	+	re
+	−	+	−	−	+	+	re
+	+	−	+	−	−	+	la dernière en théorie serait comme la 1ère
−	−	−	−	−	−	−	~~donc~~ on va mettre que des −

On a donc une matrice en variables codées (+1 ou −1)

On peut alors construire un plan expé, <u>et</u> une matrice des effets X

Avec les codes, on peut comparer les β_j entre eux et donc les trier et déterminer leur importance relative.

<u>exemple</u> : $\hat{y} = 100 + 2,5 X_1 + 7 X_2 + 9 X_3$

↳ les X étant codés, on peut définir le + important, en l'occurrence X_3 dont le coef. est 9.
↳ on devra définir si le X_1 est vraiment significatif car son coef est bas.

Module 2

On fait un criblage H_N $\quad X_2 > X_3 > \underbrace{X_5 > X_4 > X_1}_{N.S.}$

On considère k facteurs

$$\hat{y}_i = \underset{\substack{\uparrow \\ \text{calculé}}}{\beta_0} + \underbrace{\sum_1^k \beta_j X_j}_{\text{effet}} + \underbrace{\sum_{ij}^{k} \beta_{ij} X_i X_j}_{1^{er}\text{ ordre}} + \underbrace{\sum_{ijk} \beta_{ijk} X_i X_j X_k}_{2^e \text{ ordre}}$$

Avec $k=3$

$$\hat{y} = \beta_0 + \sum_1^3 \beta_j X_j + \beta_{12} X_1 X_2 + \beta_{13} X_1 X_3 + \beta_{23} X_2 X_3 + \beta_{123} X_1 X_2 X_3$$

On a bien $k =$ nb facteur $\underline{\underline{\text{et}}}$ $k' = 2^k \Rightarrow N$

$$C_5^0 = \frac{5!}{0!\,5!} = 1\, b_0$$

$$C_5^1 = \frac{5!}{1!\,4!} = 5\, b_j$$

$$C_5^2 = \frac{5!}{2!\,3!} = 10\, b_{ij} \quad (1^{er} \text{ ordre})$$

$\vdots$

$$C_5^5 = \frac{5!}{5!\,0!} = 1 \text{ interaction } 4^e \text{ ordre}$$

Exercice...

Plan expérimental

Matrice des effets

→ On rajoute les 1 de la $1^{ère}$ colonne

→ Pour remplir la colonne interaction, on multiplie les colonnes "concernées"

X_1	X_2	X_1X_2
-1	-1	$+1$
$\vdots$	$\vdots$	

Dimension de $X : [2^k, 2^k] \rightarrow$ propriété d'Hadamar

$X' : [2^k, 2^k]$

$X'X : [2^k, 2^k]$

Rappel → matrice ID : diagonale en 1 et 0 de part et d'autre

TRES

6	2	2	2
2	6	2	2
2	2	6	2
2	2	2	6

Revoir dans le poly 2A de trait^{mt} des matrices.

En + d'êtres pratiques pour les calculs matrice,
Les matrices d'Hadamar permettent :
- de limiter le nb d'expériences
- d'économiser des calculs
  (les étapes du calcul matriciel ne sont pas nécessaires)

Avec la propriété d'Hadamar, le coefficient correspond à la somme des y affectés aux signes de la colonne étudiée, ~~divisée par~~ multipliée par les valeurs de la diagonale de X'X

La croix d'interaction

Module 3 — Générateurs — Demi plan complet

- quand on fait un produit, on étudie les facteurs
- y a bcp de facteurs alors on cherche les principaux
- on fait un modèle avec ces facteurs principaux qu'on définit en 2 ou 3 niveaux
- 2 min
- on fabrique les fcats

→ pour aller + finement ds l'analyse, on va faire le modèle synergique on va édifier un plan complet qui prend en compte toutes les interactions → 2^k (k étant le nb de facteurs principaux)

sauf que si $k = 2 \longrightarrow N_p = 2^2 = 4$

$3 \longrightarrow 8$

$6 \longrightarrow 64$ ← ça fait bcp d'expériences...

Mais y a des interactions de 1^{er}, 2^e, 3^e ordre

Mais à partir du 2^e ordre, les effets sont petits voire négligeables ou encore ininterprétable

- on garde seul^{nt} les 1^{er} ordre du coup.

et encore ! elles ne sont pas toutes plausibles !

↳ Au final, on peut considérer que nous n'avons besoin que de la ½ des expériences !

soit $\dfrac{2^k}{2} = 2^{k-1}$ expériences

⚠ on garde la "k−1"
et $2^3 \neq 2^{4-1}$ car
→ 2^3 est un plan complet à 3 facteurs
→ 2^{4-1} est un plan complet à 4 facteurs dont on ne fait que la moitié des expériences
↳ ainsi 2^3 a 3 colonnes dans le plan expé alors que 2^{4-1} en a 4.

Méthode — On démarre toujours d'un modèle. Prenons :

$$\hat{y} = \beta_0 + \sum_1^4 \beta_j X_j + \beta_{12} X_1 X_2 + (\beta_{14} X_1 X_4)$$

on trouve β_{14} avec $f_1 + f_4$

$\left.\begin{matrix} k = 4 \\ k' = 7 \end{matrix}\right\} 2^{4-1}$ donc $N_p = 8$ (sur 16 expé, on en prend la moitié)

nb de facteurs + intérac° ?

On construit :

	I	X_1	X_2	X_3	X_4	X_1X_2	X_1X_4
1	+	−	−	−		+	
2	+	+	−	−		−	
3	+	−	+	−		−	
4	+	+	+	−		+	
5	+	−	−	+		+	
6	+	+	−	+		−	
7	+	−	+	+		−	
8	+	+	+	+		+	

TRES

on veut remplir ces colonnes de façon à retrouver la propriété d'Hadamar

On sait que la matrice 2^3 a la propriété d'Hadamar

Dans le tableau de cette matrice on a des colonnes alternance les signes +/−

$I \quad X_1 \quad X_2 \quad X_3 \quad X_1X_2 \quad X_1X_3 - X_2X_3 - X_1X_2X_3$ — ces alternances de signe ne sont pas encore utilisées dans notre matrice 2^{4-1}

on a déjà ces alternances de signes

On va donc utiliser une (mettons X_{123}) pour l'alternance de X_4 !
On parle de relation de définition : $\underline{X_4 = X_1X_2X_3}$

On a d'autre part le générateur d'aliases

on a $X_4 X_4 = X_1X_2X_3 X_4 \implies \underline{I = 1234}$

I	X_1	X_2	X_3	~~X_4~~	X_{12}	X_{13}	X_{23}	X_{123}
1234	234	134	124		34	24	14	4

← sur cette ligne on ne met pas les X pour simplifier l'écriture

ex : 4 car $X_1X_2X_3X_4 \times X_1X_2X_3 = X_4$

si on met $X_4 = X_1X_3$ alors $I = 134$ et

I	X_1	X_2	X_3	X_{12}	X_{13}	X_{23}	X_{123}
134	34	1234	14	234	4	[illegible]	24

Module 6

si pas d'effet d'intéraction ⟶ on fait un plan de criblage
 si on considère toutes les interactions ⟶ on fait un plan complet, synergiques
si on dit que des intéractions sont négligeables ⟶ demi plan

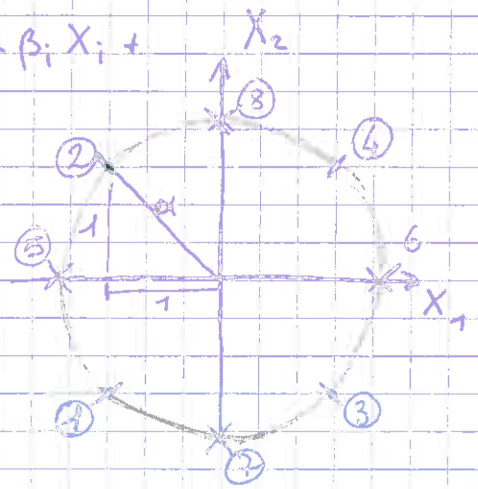
Si on (j'ai loupé l'info clef → pk au carré ?!)

⟶ on fait un plan quadratique

Exemple à 2 facteurs

$$= \beta_0 + \sum_{1}^{2} \beta_i X_i +$$

	N°	X_1	X_2
	1	−	−
N_f	2	+	−
	3	−	+
	4	+	+
	5	$-\alpha$	0
N_a	6	α	0
	7	0	$-\alpha$
	8	0	$+\alpha$
	9	0	0
N_0	10	0	0
	11	0	0
	12	0	0



$\alpha = N_f^{1/4}$

N_f, N_a et N_0 sont des genres de groupes de points (donc essais)

$N_f = 2^k$ ou 2^{k-p}

$N_a = 2 \times k$

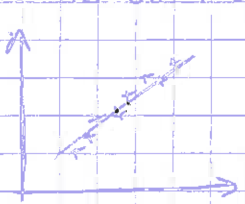
$N_0 = 4$

Quel est le niveau de facteur en variable naturelle ? $X_j \to U_j$

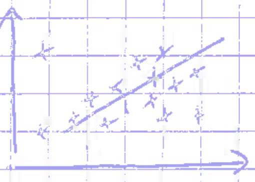
$$X_j = \frac{U_j - U_{0j}}{\Delta U_j}$$

La matrice des effets d'un plan quadratique ne possède pas la propriété d'Hadamard (donc on ne sait pas ~~te~~ calculer sans machine les autres matrices)

Valider un modèle



Cas 1 : dangereux d'approximer par une droite



Cas 2 : on peut estimer par une droite

$$SCE_y = \sum y_i^2 - n\bar{y}^2$$

$$SCE_e = SC_e = \sum e_i^2$$

$$SCE_{\hat{y}} = \sum \hat{y}_i^2 - n\hat{\bar{y}}^2 = \sum y_i^2 - n\bar{y}^2$$

selon le F_{prev}

$$F_{struct} = \frac{CM_{\hat{y}}}{CM_e} \quad \left(\text{et } F_{théo} \text{ « } F_{1-\alpha}\left(k'-1; ddl_e\right)\right)$$

ou

$$F_{struct} = \frac{CM_{\hat{y}}}{CM_{repet}} \quad (\text{et } F_{théo}$$

Facteurs de variation	SCE	ddl	CM
$\hat{y}$ (expliqué par modèle)	$SCE_{\hat{y}}$	$k'-1$	$CM_{\hat{y}}$
e (résiduelle)	SCE_e	ddl_e	CM_e
adéquation	SCE_{ad}	ddl_{ad}	CM_{ad}
répétitions	SCE_{repet}	ddl_{rep}	CM_{repet}
y totale (observée)	SCE_y	SCE_y	

$F_{1-\alpha}\left(ddl_{ad}; ddl_{rep}-1\right)$

$F_{prev} = \frac{CM_{ad}}{CM_{rep}}$

Les hypothèse : est-il prévisionnel ou pas ce modèle ? (= valable ?)

$H_0 : \frac{CM_{ad}}{CM_{repet}} = 1$ $\qquad$ $H_1 : \frac{CM_{ad}}{CM_{repet}} > 1$

$F_{prev} < F_{théo}$: le modèle est prévisionnel et $\hat{\sigma}_\varepsilon^2 = CM_e$ et $F_{théo} = F_{1-\alpha}(k'-1; ddl_e)$

$F_{prev} > F_{théo}$: on a ~~pas~~ pu prouver que le modèle n'est pas prévisionnel avec α % de risque d'erreur. et $\hat{\sigma}_\varepsilon^2 = CM_{repet}$ et $F_{théo} = F_{1-\alpha}(k'-1; r-1)$

What about le modèle structurel ?
On compare la variation des y expliqués par le modèle $\hat{y}$ à l'erreur expérimentale $\sigma_\varepsilon^2 \begin{cases} CM_e \\ CM_{repet} \end{cases}$

23/11/2012

Module 8 — Modélisation

1^er^ étape : * on postule un modèle du 1^er^ degré

$$\hat{y} = \beta_0 + \sum^{k} \beta_j X_j + \qquad 2^k \text{ ou } 2^{k-p}$$

* on valide le modèle du 1^er^ degré (test F sur les coef (optimisation

<u>NB</u> - si ce modèle n'est pas valide, on teste le modèle du second degré

$$\hat{y} = \beta_0 + \sum \beta_j X_j + \text{interaction} + \sum \beta_{jj} X_j^2$$

↳ essais en étoile $N_a = 2 \qquad + (2) N_0$

→ sur les $N_a + N_y + N_0$ réponses → valider le modèle

si le modèle est valide → test t + optimisation

si le modèle quadratique n'est pas valide, il faut refaire des expériences dans A.

Valider un modèle prévisionnel

nos hypothèses H_0 et H_1 ... patati avec $\frac{CM_{ad}}{CM_{repet}} = 1$ $\quad CM_{\hat{y}}$ → part expliqué par le modèle $\quad CM_e$ →

AV	SCE	ddl	CM	F_{calc}
expliq° par la régression $\hat{y}$	$SCE_{\hat{y}}$	$k'-1$	$CM_{\hat{y}}$	
résiduelle	SCE_e	ddl_e	CM_e	
Adéquation	SCE_{ad}	ddl_{ad}	CM_{ad}	$F_{prev} = \frac{CM_{ad}}{CM_{rep}}$ avec $F_{théo} = F_{1-\alpha}(ddl_{ad}, R-1)$
Répétition	SCE_{rep}	$R-1$	CM_{rep}	
Observé	SCE_y	$N-1$		

To know :
$$SCE_e = SCE_{ad} + SCE_{rep}$$
$$SCE_{\hat{y}} + SCE_e = SCE_y$$

Si...

$F_{prev} < F_{théo}$

conserve $H_0 \rightarrow$ model prévisionnel

$\hat{\sigma}_\varepsilon^2 = CM_e$

$F_{prev} > F_{théo}$

on a pu mettre en évidence que le modèle est non prévisionnel avec α % de risque de se tromper

$\hat{\sigma}_\varepsilon^2 = CM_{rep}$

<u>Modèle structurel</u>

$H_0 \quad \frac{CM_{\hat{y}}}{CM_e} = 1$

H_1

Plan de criblage → tri des facteurs → études des effets principaux + interac°

Modèles donné → trouver les matrices / coef et test ANOVA et t

Si

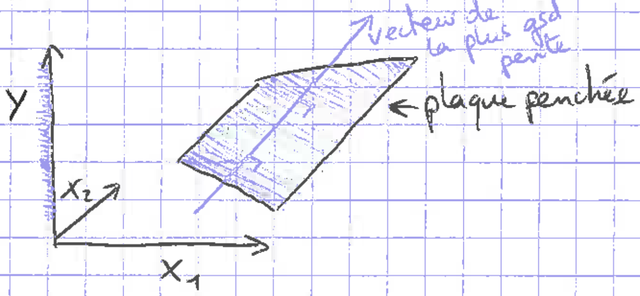
- modèle prévi → courbes isoréponse → si optimum pas atteint → trouver l'optimum en sortant du domaine sans trop faire d'expériences (économie) → méthode de la + gde pente
- modèle trooop mal connu, ou alors carrément on sait rien → méthode du simplex (pas à pas) → codage en naturel

De toute façon, faut raisonner en géométrie (coef directeurs d'un vecteur et coordonnées de triangle.

Module 9 : Méthode de la plus grande pente (optimisation)

Prérequis : connaître le modèle prévisionnel et d'avoir des courbes isoréponses

$$\hat{y} = \beta_0 + \beta_1 X_1 + \beta_2 X_2$$



On veut calculer les composantes du vecteur. Composante's

↓

1 → $b_1 \Delta U_1$

↓

2 → $b_2 \Delta U_2$

↓

j^e → $b_j \Delta U_j$

b_j étant de coef. mod. var. codée

⚠ ΔU_j est le pas de variation $\frac{U_j^{n-1} - U_j^{n+1}}{2}$

Le pas d'avancement $= \Delta_j = K \times \text{composante}_j$

$$\hat{y} = 10 + 5X_1 + 7X_2$$

$\Delta U_1 = 2$ $\quad U_{01} = 100$

$\Delta U_2 = 3$ $\quad U_{02} = 50$

	U_1	U_2
composante Δ (K = 1)	10 ; 10 $\frac{10}{7}$; =?	21 ; 2 1 3 = $\frac{21}{7}$
Départ P_0	100	50
P_0 + 1Δ	110 ; 100 + $\frac{10}{7}$	71
P_0 + 2Δ	120 ; 100 + 2 × $\frac{10}{7}$	92

En rouge si on décide d'avancer U_2 de 3 en 3

NE PAS CONFONDRE

ΔU_j et Δ_j

MRE 9

Donnée : $\hat{y} = \beta_0' + \beta_1' U_1 + \beta_2' U_2$

Recherché : $\hat{y} = \beta_0 + \beta_1 X_1 + \beta_2 X_2$

A UTILISER : $X_j = \dfrac{U_j - U_{0j}}{\Delta U_j}$

Module 10 : Simplex

Cette méthode se fait pas à pas (comme la méthode de la plus gde pente)
· n'a pas besoin de modèle (pas comme)

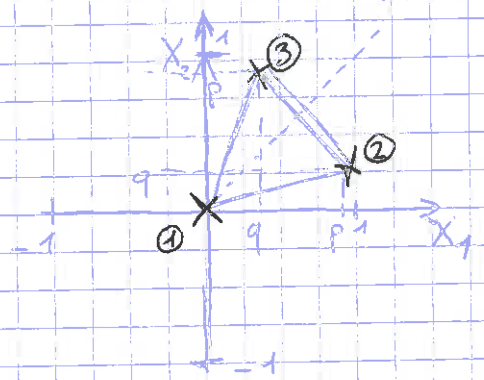
Optimiser une réponse, ça peut être maximiser (les rdts...)
ou minimiser (les pertes...)
ou encore atteindre une valeur nominale (analyse sensorielle)

On a besoin de k facteurs ds le domaine expérimental ($k \leq 6$ c'est mieux)

On réalise $k+1$ expériences

Exemple : on a 2 facteurs ($k=2$)

SIMPLEX de type 1 : on place le triangle équilatéral de façon symétrique à la 1ère bissectrice

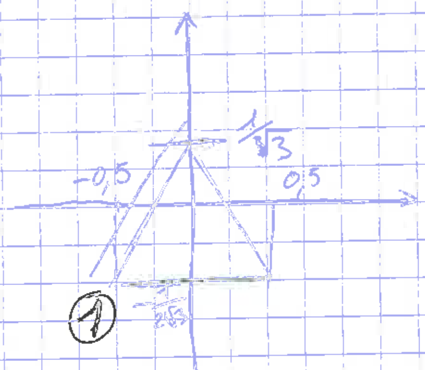


① : 1ère expérience

On construit la matrice qui contient les valeurs 0, p et q *

Les valeurs p et q sont "données"

SIMPLEX de type 2 :



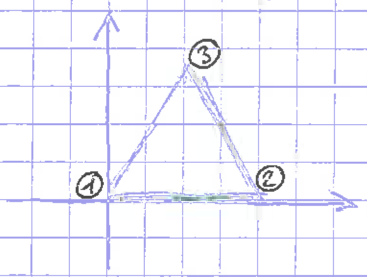
de forme

essai \ facteur	1	2	3	4
1	$-\frac{1}{2}$	$-\frac{1}{2\sqrt{3}}$		
2				
2	$\frac{1}{2}$			
3	0			
4				

essai \ facteurs →	1	2	3	4
1	0	0	0	0
2	p	q	q	q
3	q	p	q	q
4	q	q	p	q
5	q	q	q	p

* c'est de forme

SIMPEX de type 3



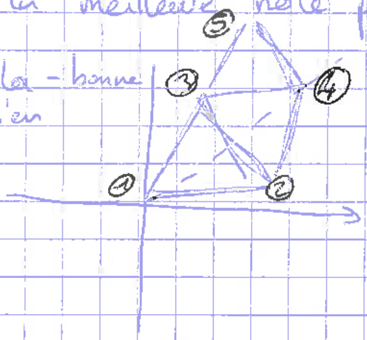
→ On doit savoir construire un simplex qu'on nous demande de faire ou encore reconnaître avec le tableau de quel simplex on parle.

Revennons à notre exemple : on veut la meilleure note possible

	X_1	X_2	~~Optimum~~ Y
1	0	0	10
2	0,966	0,259	11
3	0,259	0,966	12
123	4=1*2	?	15
234	5=2*		

on cherche la - bonne note pour s'en éloigner

bis



Dans le simplex 123 (triangle) je ~~remplace~~ place 4, ~~par~~ le symétrique de 1 par rapport à 23

puis dans le simplex 234 je place 5 le sym de 2 par rapport à 34.

⟹ Si le simplex oscille entre 2 points, on abandonne le 2^e point le moins voulu de la config. de base.

→ Si après k+1 réflexions, c'est le même sommet qui est conservé, on parle de vieillissement d'un point, soit on considère ce point optimum et on confirme ce résultat, soit on repart de ce point avec un pas de simplex plus petit

Coordonnées d'un point image

i = n° d'essai   j = n° de facteur U_{mj} = le plus mauvais point

U_{*j} = l'image du ↑

$$U_{*j} = \left(\frac{1+\nu}{k}\right) \times \left[\sum_{\substack{i=1 \\ i\neq m}}^{k+1} U_{ij}\right] - \nu U_{mj}$$

le hazard   fois   "la somme des bons"   moins   le mauvais

on remplace ν par 1 si on veut une réflexion

2 expansion

−0,5 contraction interne

0,5 contraction externe