Méthodes de criblage par discrétisation de l'espace

École Chercheur Mexico

Claude Bruchou

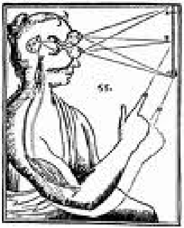
INRA, centre PACA

Giens, le 8 juin 2010

Plan

Plan

Les modèles de simulation



- incarnation d'un imaginaire moderne,
- le monde vu comme une machine formée d'un agrégat de processus interagissant,
- objectif d'emprise sur le réel,
- attrait narcissique pour le virtuel ?
- enjeu de connaissance ou outil pour justifier l'action ?

Comment fonder une critique rationnelle d'un modèle de simulation ?

S'en tenir aux avis autorisés des experts, collecter sans fin des résultats de simulations ?

La Fonction Code

Du modèle théorique à l'utilisation

- un modèle complexe
- formalisation mathématique
  - système d'équations
  - $\mathcal{G} : \Omega \subset \mathbb{R}^K \rightarrow \mathbb{R}$
    $\mathbf{x} = (x^{(1)}, .., x^{(K)}) \longmapsto \mathcal{G}(\mathbf{x})$
  - analyse théorique de $\mathcal{G}$ $\Rightarrow$ table de variation, points singuliers, optimum, inversibilité,...
  - outils : mathématicien, crayon, calcul formel
  - mais voilà, analyse difficile voire impossible !

La Fonction Code

- Formalisation informatique de $\mathcal{G}$
  - outils : analyse numérique, langage informatique
  - codage de $\mathcal{G} \Rightarrow$ **Fonction Code** (FC) $G \sim \mathcal{G}$,
    $\mathbf{x} = (x^{(1)}, .., x^{(K)}) \longmapsto y = G(\mathbf{x})$
  - arguments de G : facteurs contrôlés quantitatifs ou qualitatifs,
  - G **boite noire**
- FC **déterministe** : la répétition d'une entrée produit la même sortie,
- FC **stochastique** : la répétition d'une entrée produit des sorties différentes.

Une méthode d'analyse non intrusive

- recherche des facteurs influents de G pour un objectif ou critère $\mathcal{C}(y)$
- échantillonnage aléatoire :
  - tirage au hasard, selon des lois de probabilité, des niveaux $\mathbf{x}$ du K-uplet des facteurs,
  - Hypothèse fondamentale : **choix des gammes de variation des facteurs**,
  - facteurs considérés comme des variables aléatoires $\mathbf{X}$,
  - la loi de tirage des facteurs induit une loi de probabilité sur les sorties,
  - $\Rightarrow$ analyse statistique des sorties (analyse des moments),
  - on note y une réalisation de la FC, considérée comme une variable aléatoire Y.
- échantillonnage déterministe : niveaux fixés des facteurs,
- Illustration.

Plan

Intérêt des plans factoriels

- étude simultanée de plusieurs facteurs en entrée,
- gain en coût expérimental et en temps,
- possibilité de détecter des interactions.

Pavage de l'espace Ω

- découpage des gammes des facteurs en intervalles,
- si échantillonnage uniforme : intervalles de même amplitude
- si échantillonnage non-uniforme : amplitudes définies par les quantiles,
- R tirages uniformes indépendants dans chaque pavé.
- $\Rightarrow$ Plan complet et équilibré.

Exemple

- FC déterministe : $\Omega = [-1,1]^2 \to \mathbb{R}$, $Y = G(X^{(1)}, X^{(2)})$
- décomposition de la gamme $[-1,1]$ en I (resp. J) intervalles de même amplitude pour $X^{(1)}$ (resp. $X^{(2)}$),
- $\Rightarrow$ partition de Ω en IJ pavés (unités expérimentales),
- $I = J = 3$ et $R = 2$,
- découpage de $X^{(1)}$ (resp. $X^{(2)}$) $\Rightarrow$ facteur A (resp. B)

Exemple (suite)

simu.	classe de A	classe de B	r	sortie Y_{ijr}
1	1	1	1	9.5
2	1	1	2	10.5
3	1	2	1	20.2
4	1	2	2	18.8
5	1	3	1	30.3
6	1	3	2	27.7
7	2	1	1	8.4
8	2	1	2	9.6
9	2	2	1	29
10	2	2	2	31
11	2	3	1	41
12	2	3	2	39
13	3	1	1	1.5
14	3	1	2	2.5
15	3	2	1	31.5
16	3	2	2	28.5
17	3	3	1	41
18	3	3	2	39

moyenne $\bar{Y}_{\ldots} = 23.28$, variance $\sigma^2_Y = 180.9$

Modèle d'ANOVA à effets fixes

- $Y_{i,j,r} = \mu + A_i + B_j + AB_{i,j} + \epsilon_{i,j,r}$
  $i = 1, I, \ \ j = 1, J, \ r = 1, R$
- μ = constante,
- A_i : i^{eme} effet factoriel ; $\sum_i A_i = 0$,
- B_j : j^{eme} effet factoriel ; $\sum_j B_j = 0$,
- AB_{ij} : effet d'interaction ; $\forall j, \sum_i AB_{ij} = 0$ et $\forall i, \sum_j AB_{ij} = 0$,
- ϵ erreur *aléatoire* d'espérance nulle et de variance σ^2_Y.

Les effets

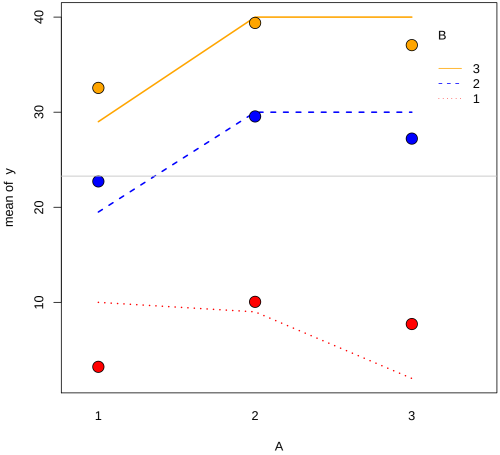
- effets principaux :
  - estimateur de A_i : $\hat{A}_i = \bar{Y}_{i..} - \bar{Y}$
  - estimateur de B_j : $\hat{B}_j = \bar{Y}_{.j.} - \bar{Y}$

	A_i	B_j
1	-3.78	-16.28
2	3.06	3.22
3	0.72	13.06

- effets d'interaction :
  - $AB_{ij} = (E(Y_{ij}) - \mu) - A_i - B_j$
  - $\Rightarrow \widehat{AB}_{ij} = \bar{Y}_{ij.} - \bar{Y}_{i..} - \bar{Y}_{.j.} + \bar{Y}$

$i \setminus j$	1	2	3
1	6.78	-3.22	-3.56
2	-1.06	0.45	0.61
3	-5.72	2.78	2.94

Interprétation graphique de l'interaction



points = modèle avec effets principaux

courbes = modèle avec effets principaux + interaction

(passent par les points d'ordonnée $\bar{Y}_{ij.}$)

Décomposition de la variabilité

$$SS_T = \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{r=1}^{R}(Y_{ijr} - \bar{Y}_{...})^2 = SS_A + SS_B + SS_{AB} + SS_\epsilon = 3075.7$$

$$SS_A = JR\sum_{i=1}^{I} A_i^2 = 6\sum_{i=1}^{3}(\bar{Y}_{i..} - \bar{Y})^2 = 144.8, \quad SS_B = IR\sum_{j=1}^{J} B_j^2 = 6\sum_{j=1}^{3}(\bar{Y}_{.j.} - \bar{Y})^2 = 2674.8$$

$$SS_{AB} = R\sum_{i=1}^{I}\sum_{j=1}^{J} AB_{ij}^2 = 2\sum_{i=1}^{3}\sum_{j=1}^{3}(\bar{Y}_{ij.} - \bar{Y})^2 = 239.6$$

$$SS_\epsilon = \sum_{i=1}^{3}\sum_{j=1}^{3}\sum_{r=1}^{2}(Y_{ijr} - \bar{Y}_{ij.})^2 = 253.11 = 16.6$$

$$n = IJR, \; V(Y) \simeq \frac{SS_T}{n} = \frac{\sum_{i=1}^{I} A_i^2}{I} + \frac{\sum_{j=1}^{J} B_j^2}{J} + \frac{\sum_{i=1}^{I}\sum_{j=1}^{J} AB_{ij}^2}{IJ} + \frac{SS_\epsilon}{n}$$

Indices de Sensibilité

- Indice = contribution de la somme des carrés d'un effet à SS_T
- si un facteur a plus de modalités que les autres, sa contribution sera plus forte
- indices utilisés en analyse de sensibilité (Saltelli) :
  - indices principaux

    facteur $X^{(1)}$: $I_1 = \dfrac{SS_A}{SS_T} = 0.047$

    facteur $X^{(2)}$: $I_2 = \dfrac{SS_B}{SS_T} = 0.87$
  - indice totaux

    facteur $X^{(1)}$: $IT_1 = \dfrac{SS_A + SS_{AB}}{SS_T} = 0.125$

    facteur $X^{(2)}$: $IT_2 = \dfrac{SS_B + SS_{AB}}{SS_T} = 0.95$

Inférence sur les Indices de Sensibilité : Bootstrap

- tirage avec remise dans chaque pavé (si $R > 1$),
- tirage avec remise dans la liste des erreurs issue d'un modèle d'anova,
- calcul des indices pour chaque échantillon boostrap,
- IC de probabilité $1 - p$ défini par les quantiles $q_{p/2}$ et $q_{1-p/2}$ de la distribution des indices.

Analyse sous R

- tirage de l'échantillon

```
tirage.r = function(PAV,  binf, bsup, Nbclass){
#------------------------------------------------------------
#  TIRAGE uniforme des coordonnées d'un point
#          dans le pavé PAV de R^K
#------------------------------------------------------------
#  binf et bsup : vecteurs des bornes inf et sup des facteurs,
#  Nbclass : nbre de classes par facteur,
#  PAV : vecteur de K numéros de classe (1 à Nbclass)
#  sortie = vecteur de K éléments
   K = length(PAV)
   bornes = matrix(NA,nrow=K, ncol=2)
   bornes[,1]= binf + (PAV-1)*(bsup-binf)/Nbclass
   bornes[,2]= binf + PAV*(bsup-binf)/Nbclass
   cc = numeric(K)
for(i in 1:K) cc[i] = runif(1,min=bornes[i,1],max=bornes[i,2])
   cc
    }
```

Analyse sous R

- table des simulations et anova

```
> K=2;Nbclass=3; r=2; binf=c(-1,-1); bsup=c(1,1);
> facteurs = expand.grid(X=1:Nbclass, Y=1:Nbclass, R=1:r)
# simu = matrice K x nrow(facteurs)
>simu = apply(facteurs[,1:K],1, tirage.r, binf, bsup, Nbclass)
> data = data.frame(X=as.factor(facteurs[,1],
                   Y=as.factor(facteurs[,2],
                   Z =  mon.modele(simu) )
 >aov.out= aov(Z ~X + Y + X:Y, data) ou aov(Z ~(X + Y)^2, data)
 > summary(aov.out)
```

	Df	Sum Sq	Mean Sq	F value	Pr($>F$)
A	2	144.78	72.39	39.294	3.573e-05
B	2	2674.78	1337.39	725.965	1.130e-10
AB	4	239.56	59.89	32.509	2.327e-05
Residuals	9	16.58	1.84		

Quid des tests de Fisher F en simulation ?

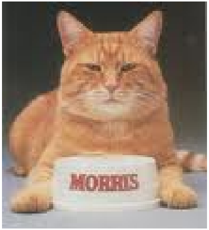
- FC déterministe, population = noeuds d'une grille
  - exploration exhaustive et modèle d'anova sans erreur $\Rightarrow$ test F dépourvu de sens.
- FC déterministe, population = Ω domaine continu
  - **échantillon aléatoire**, erreur définie par le modèle d'anova $\Rightarrow$ test F valide,
  - différence significative $\neq$ différence importante.

Bilan

- souplesse, interprétation, pistes pour un méta-modèle,
- nombreux plans dans la littérature,
- modèle linéaire limité pour certaines non-linéarités de la FC,
- plan complet souvent inabordable : 30 facteurs, 3 intervalles, $R = 1$, $\Rightarrow 3^{30} \sim 2 \times 10^{14}$ simulations. $t_{exec} = 10^{-4}s$, 1000 processeurs $\Rightarrow T_{tot} = 7.8$ mois.

Plan

Généralités



- méthode exploratoire
- FC à temps de calcul élevé
- plan OAT (One At a Time) : on fait varier un seul facteur à la fois.

Plan d'expérience standard

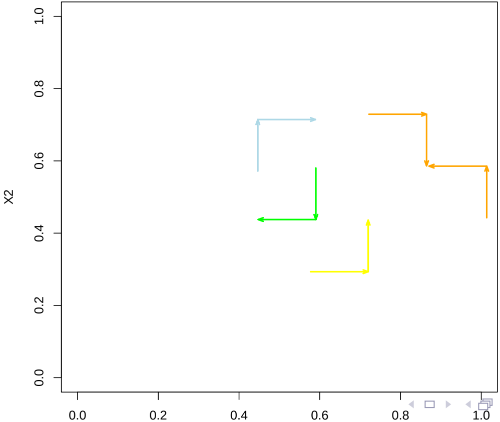
- K facteurs definis dans $[0, 1]$,
- loi d'échantillonage uniforme pour chaque facteur,
- discrétisation des facteurs en Q niveaux équiprobables,
- $\tilde{\Omega}_i = \{0, \frac{1}{Q-1}, \frac{2}{Q-1}, .., \frac{Q-1}{Q-1} = 1\} \subset [0, 1]$,
- grille $\tilde{\Omega} = \otimes_{i=1}^{K} \tilde{\Omega}_i \subset [0, 1]^K$,
- un point de base P^* est choisi au hasard sur la grille,
- accroissement de $\delta > 0$ de une ou plusieurs coordonnées de P^* $\Rightarrow P_1 = (x_1^{(1)}, ..., x_1^{(K)})$ 1er point de la trajectoire dans $\tilde{\Omega}$

Plan d'expérience standard

- P_2 2ième point de la trajectoire par variation d'une **seule** coordonnée i de P_1 :
  - soit $x_j^{(i)}$ la valeur du facteur $X^{(i)}$ pour le point j
  - $P_2 = (x_2^{(1)}, ..., x_2^{(K)}) = (x_1^{(1)}, .., x_1^{(i)} + \epsilon\delta, .., x_1^{(K)}) \in \tilde{\Omega}$
  - $\epsilon = \pm 1$ orientation aléatoire,
  - $\delta \propto \frac{1}{Q-1}$ commun à tous les facteurs.
- P_3 est obtenu à partir de P_2 en faisant varier une coordonnée i' autre que i,
- P_4 est obtenu à partir de P_3 en faisant varier une coordonnée i'' autre que i et i',
- $\Rightarrow$ trajectoire constituée de P_1, P_2, ..., P_{K+1},
- tous les facteurs n'ont varié qu'une seule fois,
- N trajectoires $\Rightarrow N(K+1)$ simulations.

Exemple

- $K = 2$ facteurs discrétisés en $Q = 7$ niveaux,
- $\delta = 1/6$, $N = 5$ trajectoires de 3 points,
- la flèche indique l'ordre des points d'une trajectoire.



Indices

- **effet élémentaire** :
  - soit une trajectoire T, $(T = 1, N)$
  - variation de la sortie Y associée à une variation du facteur i :

$$\Delta_T^{(i)} Y = \frac{G(.., x^{(i)} + \epsilon\delta, ..) - G(.., x^{(i)}, ..)}{\epsilon\delta}$$

- **moyenne** des variations de Y associées aux variations du facteur $X^{(i)}$

$$\mu_i = \frac{1}{N} \sum_{T=1}^{N} \Delta_T^{(i)} Y$$

- **Rem.** : les facteurs sont ici définis sur un cube et sont sans unité. En général, les indices sont exprimés dans l'unité de Y.

Indices

- si $\mu_i \simeq 0$: le facteur $X^{(i)}$ n'a pas d'effet ou a une variation périodique,
- la moyenne des **valeurs absolues** des variations est préférable

$$\mu^{\star}{}_i = \frac{1}{N} \sum_{T=1}^{N} |\Delta_T^{(i)} Y|$$

Exemple (suite)

T	$X^{(1)}$	$X^{(2)}$	Y	$\epsilon^{(1)}$	$\epsilon^{(2)}$	$\Delta_T^{(1)} Y$	$\Delta_T^{(2)} Y$
1	0.57	0.29	8.3				
	0.71	0.29	13.9	+1		(13.9-8.3)/.14 = 40	
	0.71	0.43	6.9		+1		(6.9-13.9)/.14=-50
2	0.57	0.57	-5.7				
	0.57	0.43	1.3		-1		-50
	0.43	0.43	-4.3	-1		40	
3	0.43	0.57	-11.3				
	0.43	0.71	-18.3		+1		50
	0.57	0.71	-12.7	+1		-45.7	
4	0.71	0.71	-7.1				
	0.85	0.71	-1.1	+1		42.85	
	0.85	0.57	5.9		-1		50
5	1.00	0.43	18.5				
	1.00	0.57	11.5		+1		-50
	0.86	0.57	5.9	-1		40	
moy $\|\Delta\|$						41.71	50
e.c.t. Δ						38.7	54.8

Rem. 1 : la paire de coordonnées qui est modifiée dans une trajectoire est indiquée dans une couleur différente selon le facteur en jeu.
Rem. 2 : $\delta = 0.14$

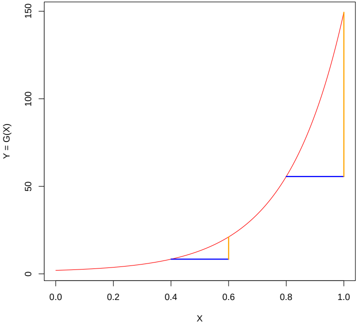
Indices

- **écart type** des variations de Y associées aux variations du facteur $X^{(i)}$

$$\sigma_i = \sqrt{\frac{1}{N}\sum_T (\Delta_T^{(i)} Y - \mu_i)^2}$$

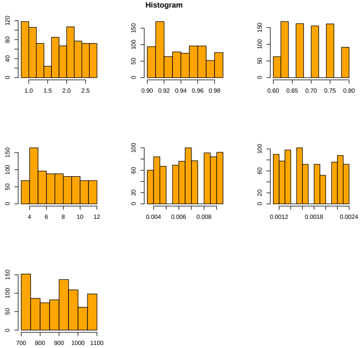
- $\sigma_i \simeq 0$, les quantités $\Delta_T^{(i)} Y$ ne sont pas influencées par les autres facteurs (i.e.pas d'interaction),
- $\sigma_i \gg 0 \Rightarrow$ interaction de X^i **ou** relation non linéaire entre $X^{(i)}$ et Y,

Une liaison non-linéaire entre la sortie de la FC et un facteur peut expliquer la variabilité des variations élémentaires.



Choix des Paramètres

- discrétisation à partir des quantiles de la loi d'échantillonage,
- si le nombre de niveaux Q **pair** et si $\delta = \frac{Q}{2(Q-1)}$, alors la distribution des coordonnées est proche de celle d'une loi uniforme
- Ex. : distribution empirique des coordonnées :



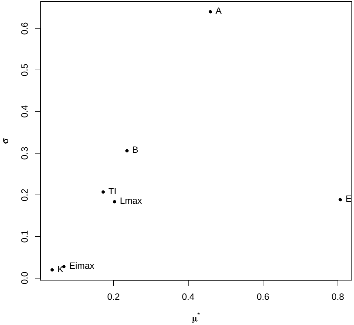
Analyse de Morris sous R

```
> Q=6
> delta = Q/(2*(Q-1))
> morris.out = morris(model = wwdm.simule,
         factors= noms , r = 30,
         design = list(type = "oat", levels = Q,
         grid.jump = delta*(Q-1), scale=F,
         binf = 0, bsup= 1)
> plot(morris.out)
```

ATTENTION : la version v1.1 de la fonction morris() du package sensitivity est à utiliser uniquement avec scale=F. De plus, les indices sont calculés correctement si les facteurs sont de **même unité**. On utilisera la gamme par défaut $[0, 1]$ avec transformation au niveau de la FC.

Interprétation du graphique de Morris

- graphique formé des points $(\mu^{\star}{}_j \times \sigma_j)$, $j = 1, K$



- points proches de l'origine : facteurs ayant un effet négligeable,
- points situés à droite sur l'axe des abscisses : facteurs ayant un effet linéaire important,
- points situés en haut à droite : facteurs ayant un effet non-linéaire ou interagissant.

Inférence sur les indices de Morris

- hypothèse d'indépendance des $|\Delta^{(i)}Y|, k = 1, N$
- intervalles de confiance (IC) bilatéraux de probabilité $1-\alpha$:
  - $\mu_i^\star \pm T_{n-1}(1-\alpha/2)\sigma_{\mu_i^\star}/\sqrt{N}$ (T_{n-1} Student)
  - $[\sqrt{\frac{(N-1)}{\chi^2(N-1,\alpha/2)}}\sigma_i, \sqrt{\frac{(N-1)}{\chi^2(N-1,1-\alpha/2)}}\sigma_i]$
- IC par bootstrap : tirage avec remise des trajectoires.
- Quel nombre de trajectoires ?

Calcul des IC de probabilité 0.95($\alpha = 0.05$) sous R

```
> IC.mu = apply(abs(x$ee),2, t.test)
> for(i in 1:7) print(IC.mu[[i]]$conf.int)
> sigma = sqrt(apply(x$ee,2,var))
> N = etude.morris$r
> for(i in 1:7) print(c(sigma[i]*(N-1)^.5/qchisq(.975,N-1)^.5,
                         sigma[i]*(N-1)^.5/qchisq(.025,N-1)^.5))
```

$\mu^\star$	Eb	Eimax	K	Lmax	A	B	TI
binf	0.73	0.065	0.033	0.19	0.38	0.15	0.15
bsup	0.80	0.08	0.042	0.28	0.64	0.27	0.2
σ							
binf	0.16	0.02	0.021	0.21	0.61	0.25	0.18
bsup	0.22	0.03	0.028	0.27	0.80	0.33	0.24

Plan

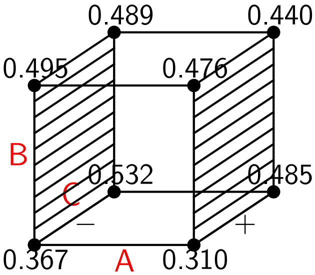
Plan complet

3 facteurs à 2 niveaux $(-1, +1)$, $N = 2^3$ simulations

MU	A	B	C	AB	AC	BC	ABC	Y
$+1$	-1	-1	-1	$+1$	$+1$	$+1$	-1	0.367
$+1$	-1	-1	$+1$	$+1$	-1	-1	$+1$	0.532
$+1$	-1	$+1$	-1	-1	$+1$	-1	$+1$	0.495
$+1$	-1	$+1$	$+1$	-1	-1	$+1$	-1	0.489
$+1$	$+1$	-1	-1	-1	-1	$+1$	$+1$	0.310
$+1$	$+1$	-1	$+1$	-1	$+1$	-1	-1	0.485
$+1$	$+1$	$+1$	-1	$+1$	-1	-1	-1	0.476
$+1$	$+1$	$+1$	$+1$	$+1$	$+1$	$+1$	$+1$	0.440

colonne AB = interaction = produit terme à terme des colonnes A et B.
L'interaction d'ordre 3 est utilisée pour estimer la variance d'erreur.

Effets factoriels



Effet de A : demie-différence des Y aux niveaux $+$ et $-$ de A.

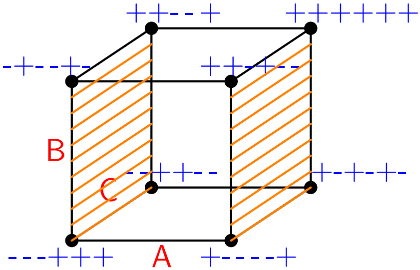
Construction

On suppose les facteurs de base A, B et C sans interactions.
Construction de 3 nouveaux facteurs.

MU	A	B	C	$D = AB$	$E = AC$	$F = BC$	ABC
$+$	$-$	$-$	$-$	$-$	$+$	$+$	$+$
$+$	$-$	$-$	$+$	$+$	$+$	$-$	$-$
$+$	$-$	$+$	$-$	$+$	$-$	$+$	$-$
$+$	$-$	$+$	$+$	$-$	$-$	$-$	$+$
$+$	$+$	$-$	$-$	$+$	$-$	$-$	$+$
$+$	$+$	$-$	$+$	$-$	$-$	$+$	$-$
$+$	$+$	$+$	$-$	$-$	$+$	$-$	$-$
$+$	$+$	$+$	$+$	$+$	$+$	$+$	$+$

- les effets de D et AB ne peuvent être estimés séparément ; ils sont **confondus** ou aliasés.
- 8 simulations au lieu de $2^6 = 64$
- plan 2^{6-3} = fraction $1/8$ du plan complet 2^6.

Effets factoriels



$$(- + - - + -) = (A = -1, B = +1, C = -1, D = -1, E = +1, F = -1)$$

Effet : demie-différence des Y aux niveaux $+$ et $-$ d'un facteur.

Modèle d'ANOVA

- notation pour un plan à 3 facteurs à 2 niveaux :
  L'espérance du vecteur des sorties Y est la somme des effets $e(.) \in \mathbb{R}$

$$\mathbb{E}(Y) = \sum_{a,b,c} A^a B^b C^c e(A^a B^b C^c)$$

  $a, b, c \in \{0, 1\}$, (ex. $A^1 B^1 C^0 = AB$ vecteur colonne)
- Exemple :
  - plan complet
    $\mathbb{E}(Y_{+,+,-}) = \mu + e(A) + e(B) - e(C) + e(AB) - e(AC) - e(BC) - e(ABC)$
  - plan fractionnaire
  - $\mathbb{E}(Y_{+,+,-,+,-,-}) =$
    $\mu + e(A) + e(B) - e(C) + e(D) - e(E) - e(F) - e(ABC)$

Modèle d'ANOVA

- effet $e(A^aB^bC^c) = 1/N < A^aB^bC^c, Y >$ (p.s. de colonnes) avec $a, b, c \in \{0, 1\}$ Ex. : $N = 8$, $e(A^1B^1C^0) = e(AB) = 1/8 < AB, Y >$.
- Effet principal de A :
$e(A) = \frac{1}{2}(\bar{Y}_{+..} - \bar{Y}_{-..}) = \frac{1}{8}[(Y_{+--} + Y_{+-+} + Y_{++-} + Y_{+++}) - (Y_{---} + Y_{--+} + Y_{-+-} + Y_{-++})] = \frac{1}{8} < A, Y >$
- Effet d'interaction (plan complet)
$e(AB) = \frac{1}{2}[e(A|B = +1) - e(A|B = -1)]$
$e(AB) = \frac{1}{2}\{\frac{1}{4}[Y_{++-} + Y_{+++} - Y_{-+-} - Y_{-++}] - \frac{1}{4}[Y_{+--} + Y_{+-+} - Y_{---} - Y_{--+}]\} = \frac{1}{8} < AB, Y >$

Résolution d'un plan fractionnaire

- un plan de résolution $\mathcal{R}$ permet d'analyser sans confusions :
  - les interactions d'ordre $(\mathcal{R}-1)/2$ si $\mathcal{R}$ est impair,
  - les interactions d'ordre $(\mathcal{R}-2)/2$ si $\mathcal{R}$ est pair.
- Ex. : le plan 2^{6-3}, de résolution $\mathcal{R}=3$, est noté 2^{6-3}_{III}.Il permet d'estimer les effets principaux sans confusions.

Analyse sous R

Génération automatique du plan (fct regular.fraction, SAS proc factex),

```
arguments de regular.fraction:
 nb facteurs, nb niveaux, r tq N = p^r unités, résolution
> plan = regular.fraction(6,2,3,3)$plan
> fich = data.frame(A=as.factor(plan[,1]),
           B=..., F=as.factor(plan[,6]),
           Y=c(.367, .532,.495, .489, .310, .485, .476, .440))
> fich.aov = aov(Y ~ A + B + C + D +E + F, fich)
> summary(fich.aov)
```

Analyse sous R

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
A	1	0.0037	0.0037	18.49	0.15
B	1	0.0053	0.0053	26.52	0.12
C	1	0.0111	0.0111	55.5	0.085
D	1	0.00016	0.00016	0.8	0.53
E	1	0.00005	0.00005	0.25	0.70
F	1	0.018	0.018	91.203	0.066
Residuals	1	0.0002	0.0002		

Plans possibles

Relation entre nombre K de facteurs et nombre N de simulations :

- si $\mathcal{R} = III$: $N \geq 1 + K$,
- si $\mathcal{R} = IV$: $N \geq 2K$,
- si $\mathcal{R} = V$: $N \geq 1 + K + K(K-1)/2$.

Nombre *maximal* κ de facteurs pour un plan de résolution V :

- de taille $N = 2^s$,
- donnant des estimateurs de variance minimale,

s	4	5	6	7	8	9
N	16	32	64	128	256	512
κ	5	6	8	11	17	≥ 23

Bilan

PROBLEME	METHODE
Très nombreux facteurs, sélectionner les plus influents	Plans de screening (résolution III ou IV)
Etudier l'influence simultanée de nombreux facteurs avec peu d'observations, détecter les principales interactions	Plans factoriels à 2 niveaux complets ou fractionnaires, résolution $\geq$ V
Etude plus détaillée de facteurs quantitatifs	Plans factoriels 3^n ou 4^n, cf surfaces de réponses

Plan

- Jacques Goupy, Lee Creighton, **Introduction aux plans d'expériences**, (2009), Dunod
- Saltelli A. et al., **Sensitivity Analysis** (2000), Wiley
- Saltelli A. et al., **Global Sensitivity analysis The primer** (2008), Wiley
- F. Fabre, C. Bruchou, A. Palloix, B. Moury, Key determinants of resistance durability to plant viruses : insights from a model within-and between-host dynamics, Virus Research 141 (2009) 140-149.
- C. Bruchou, **Chapitre 3 du manuel EC Mexico** (2010).

Plan

Mettez en oeuvre les 3 méthodes sur le modèle wwdm.

Merci de votre attention ...

En route avec Morris !

Licence

Parler en français ? C'est ringuard. It is always better in english...