


université  
PARIS-SACLAY

FACULTÉ  
DES SCIENCES  
D'ORSAY

## Analyse des entrées et sorties d'un modèle: précision et sensibilité

Marc Girondot, Université Paris Saclay  
marc.girondot@universite-paris-saclay.fr


1



## Plan

- 1/ Les deux composantes d'un modèle
- 2/ Incertitude sur des valeurs empiriques
  - L'écart-type, l'erreur standard et autres
- 3/ Incertitude sur les paramètres du modèle
  - L'erreur standard et la distribution par MCMC
- 4/ Impact des paramètres sur les sorties d'un modèle
  - L'élasticité: « One at the time »
  - La sensibilité par décomposition de la variance


2



## Notions à retenir

- Robustesse
- Ecart-type, erreur standard, quantiles
- Matrice hessienne
- MCMC par méthode bayésienne et algorithme de Metropolis-Hastings
- Elasticité
- Sensibilité, Indice de Sobol

3



Partie 3

## INCERTITUDE SUR LES PARAMÈTRES DU MODÈLE

4

## Maximum de vraisemblance

- Etant donné un échantillon observé  $x$  et une fonction de vraisemblance  $f$ , la vraisemblance quantifie la probabilité que les observations proviennent effectivement d'un échantillon (théorique) de la loi  $f$ .
- La fonction de vraisemblance, notée  $L(x_1, \dots, x_n \mid \theta_1, \dots, \theta_k)$  est une fonction de probabilités conditionnelles qui décrit les valeurs  $x_i$  (observations) d'une loi statistique en fonction des paramètres  $\theta_j$ .
- Estimer un paramètre  $x_i$  par la méthode du maximum de vraisemblance, c'est proposer comme valeur de ce paramètre celle qui rend maximale la vraisemblance, à savoir la probabilité d'observer les données comme réalisation d'un échantillon de la loi  $f$ .

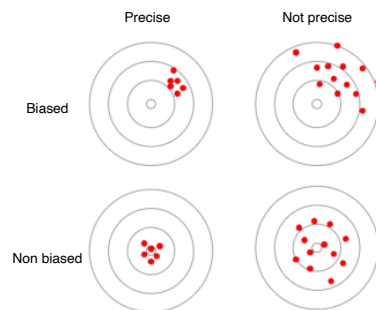
5

## L'estimation au ML

- Les paramètres  $\theta_1, \dots, \theta_k$  qui maximisent la vraisemblance  $L(x_1, \dots, x_n \mid \theta_1, \dots, \theta_k)$  correspondent à une estimation ponctuelle (*point estimation*).
- Ils représentent la combinaison qui maximise la probabilité d'observer les données comme réalisation d'un échantillon de la loi  $f$  mais ils ne nous donnent pas d'information sur la précision à laquelle on les connaît.

6

## Toujours bien de se rappeler...



- Il est difficile de déterminer qu'un estimateur n'est pas précis ou biaisé car cela implique de connaître la vraie valeur;
- Il y a souvent un compromis entre précision et biais. On ne sait pas forcément ce qui devrait être privilégié: estimateur précis ou non biaisé?

7

## Estimer la précision d'un paramètre

- Estimer la précision à laquelle est connue un paramètre est très important car cela permet de calculer son intervalle de confiance.
- On pourra alors conclure par exemple s'il est différent ou non de 0, ou bien si deux conditions expérimentales donnent des résultats significativement ou non ou si deux populations sont significativement différentes pour le caractère mesuré.

8

## L'ERREUR-STANDARD

9

## Exemple avec une distribution binomiale

- La loi binomiale modélise un évènement répété où il y a deux sorties possibles (A et B). Cette loi a un paramètre,  $p$  qui désigne la probabilité d'avoir un A (et donc  $q=1-p$  est la probabilité d'avoir un B).
- La probabilité d'une réalisation particulière d'obtention de  $k$  objets A sur un total de  $n$  s'écrit:

$$P(x = k) = \binom{n}{k} p^k (1-p)^{n-k}$$

$$\binom{n}{k} = C_n^k = \frac{n!}{k!(n-k)!}$$

$$(n-1)! = \Gamma(n)$$

10

## Autres lois possible

- On va considérer ici que le modèle binomial n'est pas une variable en soi.
- Notez quand même qu'il existe d'autres lois possibles pour décrire ces données, par exemple la loi bêta-binomiale lorsque le paramètre  $p$  de la binomiale est aléatoire et donné par une loi bêta:
- [https://fr.wikipedia.org/wiki/Loi\\_bêta-binomiale](https://fr.wikipedia.org/wiki/Loi_bêta-binomiale)

11

## Exemple avec une distribution binomiale

```
> k <- 17 # Nombre de A
> n <- 50 # Nombre total de tirage
> p <- 0.1 # probabilité de tirer un A
> choose(n, k)*p^k*(1-p)^(n-k)
[1] 3.043151e-06
> dbinom(x=k, size=50, prob = p, log=FALSE)
[1] 3.043151e-06
> dbinom(x=k, size=50, prob = p, log=TRUE) # même chose en log
[1] -12.70262
> -dbinom(x=k, size=50, prob = p, log=TRUE) # même chose en -log
[1] 12.70262
```

12

## Profil de vraisemblance

Quelle est la probabilité d'observer  $k=17/n=50$  dans un modèle binomial ( $p, 1-p$ ).

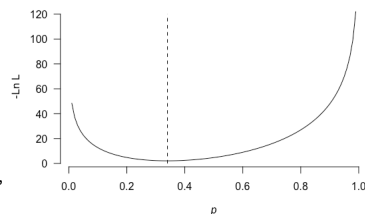
Le modèle est défini à la fois par la loi binomiale et la valeur de  $p$ .

```
p <- seq(from=0.01, to=0.99, by=0.01)
```

```
LnL <- dbinom(x=k, size=50,
  prob = p, log=TRUE)
```

```
plot(p, LnL,
  xlab=expression(italic("p")),
  ylab="-Ln L",
  bty="n", las=1,
  type="l")
```

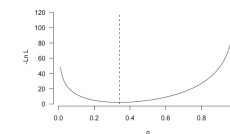
```
segments(x0=p[which.min(LnL)],
  y0=0, y1=120, lty=2)
```



13

## Maximum de vraisemblance

```
p <- seq(from=0.01, to=0.99, by=0.01)
LnL <- dbinom(x=k, size=50, prob = p, log=TRUE)
plot(p, LnL, xlab=expression(italic("p")), ylab="-Ln L",
  bty="n", las=1, type="l")
segments(x0=p[which.min(LnL)], y0=0, y1=120, lty=2)
```



- On va utiliser un algorithme qui va rechercher les valeurs des paramètres ( $p, q=1-p$ ) qui maximisent la vraisemblance des observations dans une loi binomiale. Ceci est fait avec la fonction `optim()` de R.

```
> negdbinom <- function(x, size, prob, log = FALSE) return(-dbinom(x, size,
  prob, log))
```

```
> o <- optim(par = c(prob=0.1), fn=negdbinom, size=50, x=k, log=TRUE,
  method = "Brent", lower=0, upper=1)
```

```
> o$par # C'est la valeur de p qui maximise la vraisemblance
[1] 0.34
```

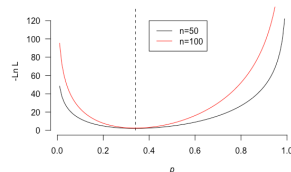
14

## Le profil de vraisemblance dépend de la taille de l'échantillon

```
k <- 17 # Nombre de A
n <- 50 # Nombre total de tirage
p <- seq(from=0.01, to=0.99, by=0.01)
LnL <- dbinom(x=k, size=n, prob = p, log=TRUE)
plot(p, LnL, xlab=expression(italic("p")), ylab="-Ln L",
  ylim=c(0, 130),
  bty="n", las=1, type="l")
segments(x0=p[which.min(LnL)], y0=0, y1=1300, lty=2)
```

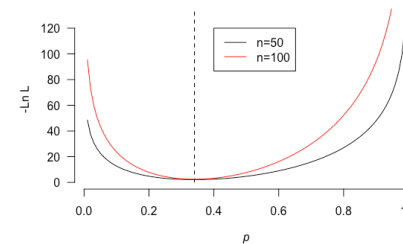
```
k <- 34 # Nombre de A
n <- 100 # Nombre total de tirage
p <- seq(from=0.01, to=0.99, by=0.01)
LnL <- dbinom(x=k, size=n, prob = p, log=TRUE)
lines(p, LnL, col="red")
```

```
legend(x=0.4, y=120, legend = c("n=50", "n=100"),
  lty=1, col=c("black", "red"))
```



15

## Le profil de vraisemblance dépend de la taille de l'échantillon



Plus l'échantillon est grand, plus le profil de vraisemblance est aigu au niveau du point ML.

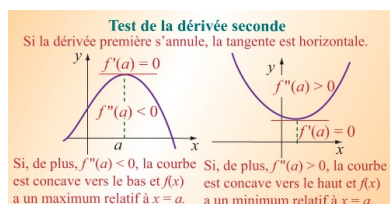
Cela signifie que quand on change la valeur du paramètre  $p$ , la vraisemblance se dégrade plus rapidement lorsque la taille de l'échantillon est importante que lorsque la taille de l'échantillon est plus faible.

On va utiliser une mesure de cette courbure au point ML pour décrire la précision avec laquelle on connaît le paramètre  $p$ .

16

## Mesurer le changement de la vraisemblance au point de ML ?

- Petit rappel de maths:
  - La dérivé première en un point mesure la pente de la tangente en ce point; elle est nulle pour un extremum
  - La dérivée seconde en un extremum mesure la courbure en cet extremum.



17

## Erreur standard de $p$

- L'erreur standard de  $p$  est alors la racine carrée de l'inverse de la dérivée seconde de la vraisemblance  $L$  au point ML:

$$SE(p) = \sqrt{\frac{1}{\frac{d^2 L}{d p^2}}}$$

18

## Effet de la taille de l'effectif sur SE

```
> k <- 17 # Nombre de A
> n <- 50 # Nombre total de tirage
> o <- optim(par = c(prob=0.1), fn=negdbinom, size=n, x=k,
log=TRUE, method = "Brent", lower=0, upper=1, hessian = TRUE)
> sqrt(1/o$hessian)
0.0669921
```

```
> k <- 34 # Nombre de A
> n <- 100 # Nombre total de tirage
> o <- optim(par = c(prob=0.1), fn=negdbinom, size=n, x=k,
log=TRUE, method = "Brent", lower=0, upper=1, hessian = TRUE)
> sqrt(1/o$hessian)
0.04737057
```

19

Ludwig Otto Hesse (22 avril 1811 à Königsberg, Prusse –  
4 août 1874 à Munich, Allemagne)



## Extension dans un système à $n$ variables

- La matrice hessienne (ou simplement la hessienne), Hessian matrix en anglais, d'une fonction numérique  $f$  est la matrice carrée, notée  $H(f)$ , de ses dérivées partielles secondes.

$$H(f) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{bmatrix}$$

En pratique, la fonction de vraisemblance n'est souvent pas dérivable et on utilise des approximations numériques de cette matrice.

20

## Extension dans un système à $n$ variables

- L'utilisation de la matrice hessienne permet de prendre en compte les covariances entre paramètres qui souvent ne sont pas nulles.

$$H(f) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{bmatrix}$$

21

## Extension dans un système à $n$ variables

- Par exemple, ici, les covariances sont supposées nulles :

◦ Girondot, M. (1999) Statistical description of temperature-dependent sex determination using maximum likelihood. *Evolutionary Ecology Research*, 1, 479-486.

◦ Godfrey, M.H., Delmas, V. & Girondot, M. (2003) Assessment of patterns of temperature-dependent sex determination using maximum likelihood model selection. *Ecoscience*, 10, 265-272.

- Alors qu'en réalité elles ne le sont pas - corrigé ici:

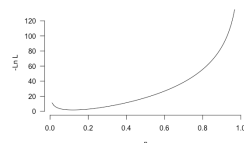
◦ Hulin, V., Delmas, V., Girondot, M., Godfrey, M.H. & Guillon, J.-M. (2009) Temperature-dependent sex determination and global change: Are some species at greater risk? *Oecologia*, 160, 493-506.

22

## Problèmes de cette méthode

- Si le profil de vraisemblance est plat autour du point ML, la dérivée seconde est nulle et SE ne peut pas être estimé.
  - Un pis-aller pour s'en sortir est de postuler que  $1/0=0$ . C'est ce que l'on appelle une matrice inversée généralisée (par exemple Moore-Penrose generalized inverse).
- La méthode suppose que le profil de vraisemblance est symétrique de part et d'autre du point ML.

```
k <- 6
n <- 50
p <- seq(from=0.01, to=0.99, by=0.01)
LnL <- -dbinom(x=k, size=n, prob = p, log=TRUE)
plot(p, LnL, xlab=expression(italic("p")), ylab="-Ln L",
      ylim=c(0, 130),
      bty="n", las=1, type="l")
```




23

## LE MCMC BAYÉSIEEN


24



Andrei A. Markov (1866-1922) 


## Chaîne de Markov

- Un processus de Markov est un processus stochastique possédant la propriété de Markov :
  - l'information utile pour la prédiction du futur est contenue dans l'état présent du processus.



PSPPSSPPPPSSSSPPSPSSSSPPSPSPPPSSPPSPSSSPSPSPPPSSPPSSSP...

25

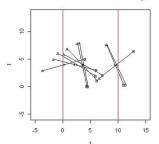
Monte Carlo est un quartier de Monaco, une ville-État. 

## Monte Carlo


- Le terme méthode de Monte-Carlo désigne une famille de méthodes algorithmiques visant à calculer une valeur numérique approchée en utilisant des procédés aléatoires, c'est-à-dire des techniques probabilistes.

L'aiguille de Buffon est une expérience de probabilité proposée en 1733 par Georges-Louis Leclerc de Buffon, au XVIIIe siècle. Cette expérience fournit une approximation du nombre  $\pi$ .

Simulation avec R de 10 lancers de baguette

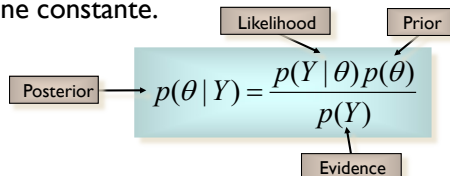


26

Thomas Bayes (1702-1761) 

## Règle de Bayes

- Permet de calculer la probabilité d'un événement (le posterior) en prenant en compte des données  $p(Y/q)$  et une information que l'on a sur cet événement (le prior).
- Si on n'a pas d'information sur les priors, cela revient à prendre la vraisemblance puisque  $p(Y)$  est une constante.



$$p(\theta | Y) = \frac{p(Y | \theta) p(\theta)}{p(Y)}$$

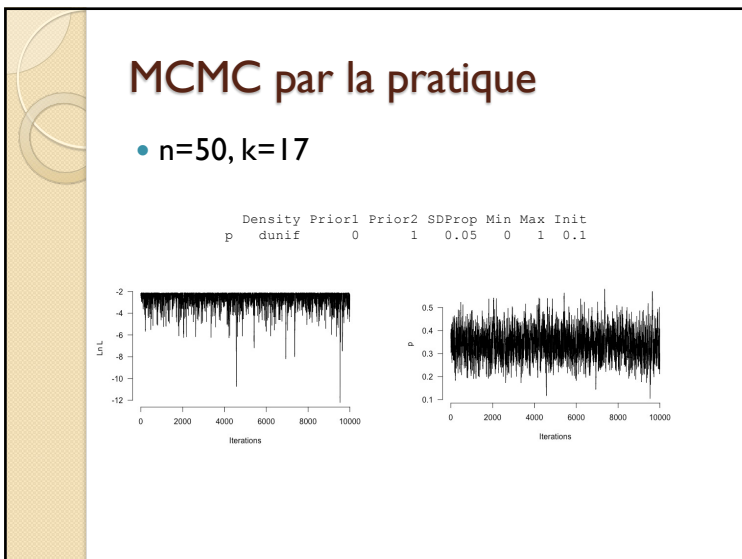
27

## Revenons à notre exemple

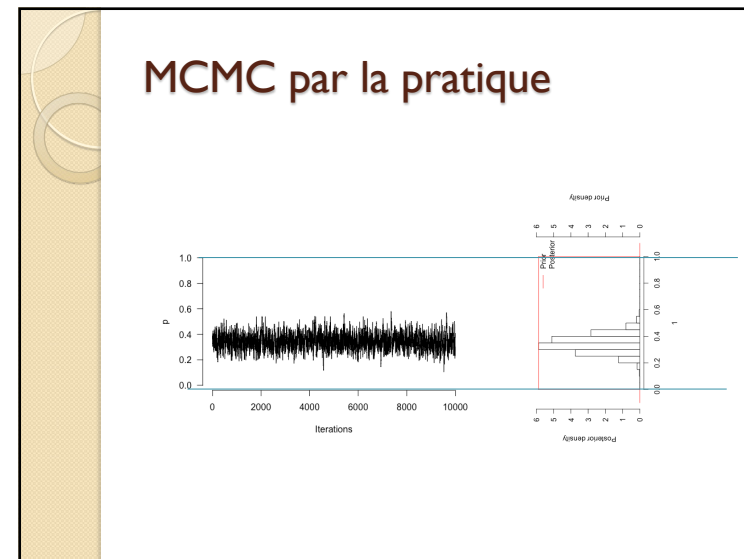
- On cherche la distribution  $p$ . On a les données  $n=50$  et  $k=17$ . On doit choisir un prior pour  $p$ . qu'a-t-on comme information sur  $p$  ? Rien. Donc on choisira un prior de  $p$  dans une loi uniforme comprise dans  $]0, 1[$ .
- On commence une chaîne de Markov avec une valeur  $p_0=0.1$  par exemple.
- On calcule la vraisemblance  $L_0$ . On tire au hasard une valeur  $p_1$  (le processus de Monte-Carlo). On calcule la vraisemblance  $L_1$  avec cette valeur. Si  $L_1$  est meilleure que  $L_0$ , on la garde sinon on peut quand même la garder avec une probabilité dépendant de sa probabilité *posterior* et de celle du point précédent: si elle est vraiment très peu probable, on a peu de chance de la garder mais ce n'est jamais nul quand même (algorithme de Metropolis-Hastings). Cela génère une chaîne de Markov.

Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H. & Teller, E. (1953) Equations of state calculations by fast computing machines. Journal of Chemical Physics, 21, 1087-1092.  
Hastings, W.K. (1970) Monte Carlo sampling methods using Markov chains and their applications. Biometrika, 57, 97-109.

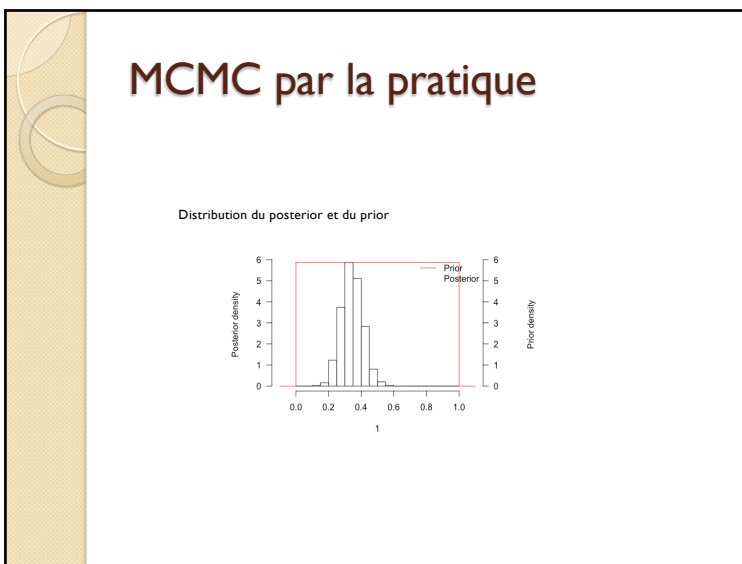
28



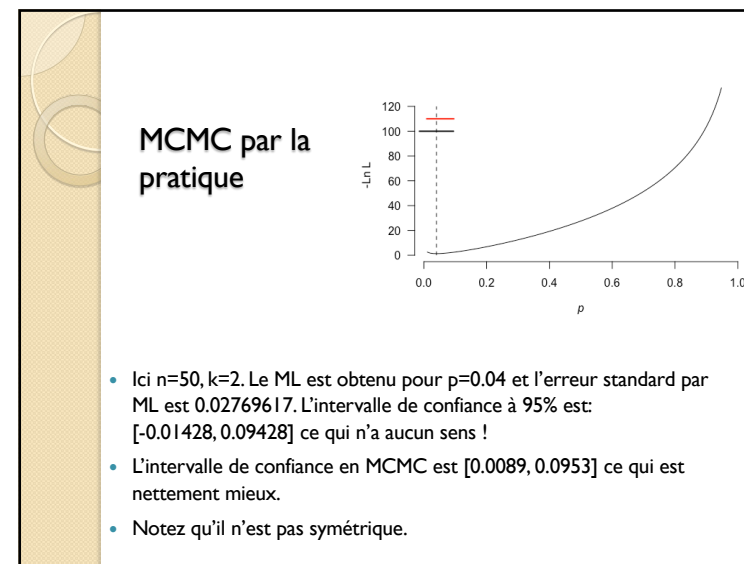
29



30



31



32



Partie 4

## ANALYSE DE SENSIBILITÉ SUR LES SORTIES D'UN MODÈLE

33

### Introduction

- Au cours de l'élaboration, de la construction ou de l'utilisation d'un modèle mathématique, l'analyse de sensibilité peut s'avérer être un outil précieux.
- En effet, en étudiant comment la réponse du modèle réagit aux variations de ses variables d'entrée, l'analyse de sensibilité permet de répondre à un certain nombre de questions.
  - 1. Le modèle est-il bien fidèle au phénomène/processus modélisé ?
  - 2. Quelles sont les variables qui contribuent le plus à la variabilité de la réponse du modèle ?
  - 3. Quelles sont au contraire les variables les moins influentes ?
  - 4. Quelles variables, ou quels groupes de variables, interagissent avec quelles (quels) autres ?

Jacques J. (2001) Pratique de l'analyse de sensibilité : comment évaluer l'impact des entrées aléatoires sur la sortie d'un modèle mathématique. *PUB. IRMA*, 71 (III), 1-19.

34

### Pourquoi une analyse de sensibilité ?

- 1. Le modèle est-il bien fidèle au phénomène/processus modélisé ?
  - En effet, si l'analyse exhibe une forte influence d'une variable d'entrée habituellement connue comme non influente, il sera nécessaire de remettre en cause la qualité du modèle ou (et) la véracité de nos connaissances sur l'impact réel des variables d'entrée.

35

### Pourquoi une analyse de sensibilité ?

- 2. Quelles sont les variables qui contribuent le plus à la variabilité de la réponse du modèle ?
  - Si cette variabilité est synonyme d'imprécision sur la valeur prédite de la sortie, il sera alors possible d'améliorer la qualité de la réponse du modèle à moindre coût en concentrant les efforts sur la réduction des variabilités des entrées les plus influentes.
  - Mais cela n'est pas toujours possible, notamment lorsque la variabilité d'une variable d'entrée est intrinsèque à la nature de la variable et non due à un manque d'information ou à des imprécisions de mesures.

36

## Pourquoi une analyse de sensibilité ?

- 3. Quelles sont au contraire les variables les moins influentes ?
  - Il sera possible de les considérer comme des paramètres déterministes, en les fixant par exemple à leur espérance, et obtenir ainsi un modèle plus léger avec moins de variables d'entrée.

37

## Pourquoi une analyse de sensibilité ?

- 4. Quelles variables, ou quels groupes de variables, interagissent avec quelles (quels) autres ?
  - L'analyse de sensibilité peut permettre de mieux appréhender et comprendre le phénomène modélisé, en éclairant les relations entre les variables d'entrée.

38

## Contexte général

- Elasticité et sensibilité sont deux termes utilisés pour décrire la relation qui existe entre une ou des variables d'entrée d'un modèle et la ou les sorties de ce modèle.
- Les termes sont utilisés par différents auteurs pour des concepts parfois différents.

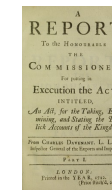
39

## Origine des concepts

- À la fin du XVIIe siècle, en particulier, Gregory King, puis Charles D'Avenant notent qu'une baisse de l'offre de blé conduit à un renchérissement du prix de cette denrée qui n'est pas en proportion de la baisse.

Gregory King (né le 15 décembre 1648 - mort le 29 août 1712) est un haut fonctionnaire britannique et l'un des premiers grands statisticiens économiques du monde moderne.

Charles d'Avenant (1656-1714), économiste anglais



40

Analyse de type "One at a time" (OAT)

## ELASTICITÉ

41

## Elasticité

- L'élasticité est le coefficient de variation d'une variable  $y$  par rapport à la variation d'une variable  $x$ .
- Le pourcentage de variation de  $y$  se notant  $\frac{\Delta y}{y}$  et celui de  $x$  se notant  $\frac{\Delta x}{x}$  on obtient, pour  $x$  et  $y$  non nuls, la formule suivante :

$$\varepsilon(y, x) = \frac{\frac{\Delta(y)}{y}}{\frac{\Delta(x)}{x}} = \frac{x}{y} \cdot \frac{\Delta(y)}{\Delta(x)}$$

42

## Elasticité

- Soit un modèle avec  $k$  paramètres. On fait varier chacun des  $k$  paramètres de  $\pm p\%$  (en général 1%) et on regarde de combien varie la sortie du modèle.
- On fait alors le rapport des taux de variation qui est l'élasticité, une valeur sans unité.

43

## Exemple avec le modèle logistique

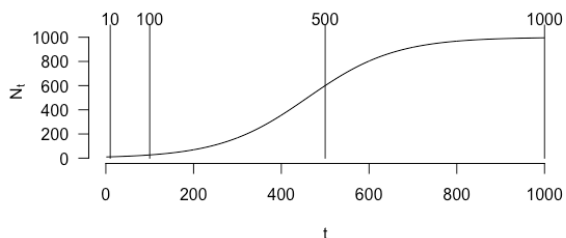
$$N_t = \frac{N_0 K}{N_0 + e^{-rt} (K - N_0)}$$

- On va chercher à évaluer l'impact de  $K$ ,  $r$  et  $N_0$  sur  $N_{10}$ ,  $N_{100}$ ,  $N_{500}$  et  $N_{1000}$  avec  $K=1000$ ,  $r=0,01$  et  $N_0=10$ .

44

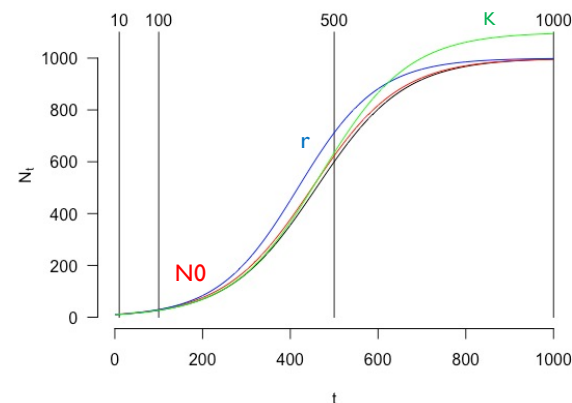
## Etat initial

```
> logistique <- function(t, N0, r, K) {(N0*K)/(N0+exp(-r*t)*(K-N0))}
> plot(x = c(1:1000), y=logistique(1:1000, N0=10, r=0.01, K=1000), type="l",
      bty="n", las=1, xlab="t", ylab=bquote("N"[t]), ylim=c(0, 1100))
> segments(x0=c(10, 100, 500, 1000), y0=0, y1=1100)
> par(xpd=TRUE); text(x=c(10, 100, 500, 1000), y=1150,
                    labels = c("10", "100", "500", "1000"))
```



45

## Effet des paramètres +10%



46

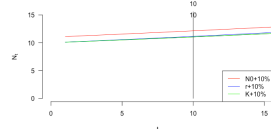
## Elasticité pour $t=10$ avec +1%

```
> N10Ini <- logistique(10, N0=10, r=0.01, K=1000)
> N10_N0_1pc <- logistique(10, N0=10+10*0.01, r=0.01, K=1000)
> Dyy <- (N10_N0_1pc-N10Ini)/N10Ini
> Dyy/0.01
[1] 0.9989389
> N10_r_1pc <- logistique(10, N0=10, r=0.01+0.01*0.01, K=1000)
> Dyy <- (N10_r_1pc-N10Ini)/N10Ini
> Dyy/0.01
[1] 0.09894436
> N10_K_1pc <- logistique(10, N0=10, r=0.01, K=1000+0.01*1000)
> Dyy <- (N10_K_1pc-N10Ini)/N10Ini
> Dyy/0.01
[1] 0.001040213
```

N<sub>0</sub>

r

K



47

## Synthèse

- Si l'élasticité  $>0$ , cela signifie qu'une augmentation sur le paramètre induit une augmentation sur la sortie
- Si l'élasticité  $<0$ , cela signifie qu'une augmentation sur le paramètre induit une baisse sur la sortie

48

## Synthèse, si l'élasticité est $> 0$

- Si l'élasticité = 0, cela signifie qu'1% d'augmentation sur le paramètre n'induit aucun changement sur la sortie
- Si l'élasticité est comprise entre 0 et 1, cela signifie qu'1% d'augmentation conduit à une augmentation de moins de 1% en sortie.
- Si l'élasticité = 1, cela signifie qu'1% d'augmentation sur le paramètre induit 1% d'augmentation sur la sortie
- Si l'élasticité est supérieure à 1, cela signifie qu'1% d'augmentation conduit à une augmentation de plus de 1% en sortie.

49

## Elasticité pour $t=100$ avec +1%

```

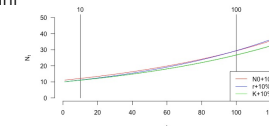
> N100Ini <- logistique(100, N0=10, r=0.01, K=1000)
> N100_N0_1pc <- logistique(100, N0=10+10*0.01, r=0.01, K=1000)
> Dyy <- (N100_N0_1pc-N100Ini)/N100Ini
> (eN0_t100 <- Dyy/0.01)
[1] 0.9829414
> N100_r_1pc <- logistique(100, N0=10, r=0.01+0.01*0.01, K=1000)
> Dyy <- (N100_r_1pc-N100Ini)/N100Ini
> (er_t100 <- Dyy/0.01)
[1] 0.9778964
> N100_K_1pc <- logistique(100, N0=10, r=0.01, K=1000+0.01*1000)
> Dyy <- (N100_K_1pc-N100Ini)/N100Ini
> (eK_t100 <- Dyy/0.01)
[1] 0.0167281

```

N<sub>0</sub>

r

K



50

## Elasticité pour $t=500$ avec +1%

```

> N500Ini <- logistique(500, N0=10, r=0.01, K=1000)
> N500_N0_1pc <- logistique(500, N0=10+10*0.01, r=0.01, K=1000)
> Dyy <- (N500_N0_1pc-N500Ini)/N500Ini
> (eN0_t500 <- Dyy/0.01)
[1] 0.4017883
> N500_r_1pc <- logistique(500, N0=10, r=0.01+0.01*0.01, K=1000)
> Dyy <- (N500_r_1pc-N500Ini)/N500Ini
> (er_t500 <- Dyy/0.01)
[1] 1.99035
> N500_K_1pc <- logistique(500, N0=10, r=0.01, K=1000+0.01*1000)
> Dyy <- (N500_K_1pc-N500Ini)/N500Ini
> (eK_t500 <- Dyy/0.01)
[1] 0.5934193

```

N<sub>0</sub>

r

K

51

## Elasticité pour $t=1000$ avec +1%

```

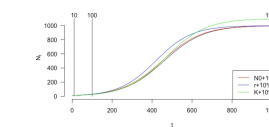
> N1000Ini <- logistique(1000, N0=10, r=0.01, K=1000)
> N1000_N0_1pc <- logistique(1000, N0=10+10*0.01, r=0.01, K=1000)
> Dyy <- (N1000_N0_1pc-N1000Ini)/N1000Ini
> (eN0_t1000 <- Dyy/0.01)
[1] 0.00447513
> N1000_r_1pc <- logistique(1000, N0=10, r=0.01+0.01*0.01, K=1000)
> Dyy <- (N1000_r_1pc-N1000Ini)/N1000Ini
> (er_t1000 <- Dyy/0.01)
[1] 0.04259847
> N1000_K_1pc <- logistique(1000, N0=10, r=0.01, K=1000+0.01*1000)
> Dyy <- (N1000_K_1pc-N1000Ini)/N1000Ini
> (eK_t1000 <- Dyy/0.01)
[1] 0.9954353

```

N<sub>0</sub>

r

K



52

## Choix de la déviation

- Le choix de 1% n'est pas neutre:

```
> N500_r_1pc <- logistique(500, N0=10, r=0.01+0.01*0.01,
K=1000)
```

```
> Dyy <- (N500_r_1pc-N500Ini)/N500Ini
```

```
> (er_t500 <- Dyy/0.01)
```

```
[1] 1.99035
```

```
> N500_r_10pc <- logistique(500, N0=10, r=0.01+0.1*0.01,
K=1000)
```

```
> Dyy <- (N500_r_10pc-N500Ini)/N500Ini
```

```
> Dyy/0.1
```

```
[1] 1.868633
```

53

## Symétrie de l'effet

- L'effet n'est pas strictement symétrique:

```
> N500_r_10pc <- logistique(500, N0=10, r=0.01+0.1*0.01,
K=1000)
```

```
> Dyy <- (N500_r_10pc-N500Ini)/N500Ini
```

```
> Dyy/0.1
```

```
[1] 1.868633
```

```
> N500_r_10pc <- logistique(500, N0=10, r=0.01-0.1*0.01,
K=1000)
```

```
> Dyy <- (N500_r_10pc-N500Ini)/N500Ini
```

```
> Dyy/-0.1
```

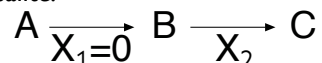
```
[1] 2.060843
```

On pourra prendre l'élasticité moyenne pour  $\pm x\%$

54

## Autres problèmes (1)

- Comment prendre en compte les interactions entre paramètres ?
  - Exemple: Le changement d'un paramètre ne change pas la sortie du modèle mais quand il change alors qu'un autre change aussi, l'effet est très important.
  - On trouvera cet effet lorsqu'il y a des chaînes de causalité:



- Changer  $X_2$  n'aura aucun effet car B est absent, mais changer  $X_1$  et  $X_2$  pourra avoir un effet très fort.

55

## Autres problèmes (2)

- Comment prendre en compte l'incertitude sur les paramètres ?
  - Cette méthode calcule l'incertitude en un point mais ce point est lui-même souvent mal connu et on risque d'explorer l'élasticité dans une zone de l'espace des paramètres qui ne correspond pas à la réalité.

56

## Gérer l'incertitude en OAT

- Morris (1991) propose d'explorer l'espace des paramètres toujours dans une stratégie OAT pour y détecter trois comportements de variables:
  - Variables sans effet sur la sortie
  - Variables avec effet fort sur la sortie
  - Variables avec effet non-linéaire

M.D. Morris. Factorial sampling plans for preliminary computational experiments. *Technometrics*, 33:161-174, 1991.

Mais en pratique ça ne résout pas les problèmes décrits précédemment.

57

## Conclusion

- C'est souvent la première analyse que l'on fait car elle est très peu coûteuse en temps de calcul.
- Mais il ne faut pas s'arrêter là car on risque de passer à côté d'effets importants ou intéressants.

58

## SENSIBILITÉ BASÉE SUR LA DÉCOMPOSITION DE LA VARIANCE

**Sobol I.M.** (1990) Sensitivity estimates for nonlinear mathematical models. *Matematicheskoe Modelirovanie*, 2, 112-118.

**Sobol I.M.** (1993) Sensitivity estimates for nonlinear mathematical models. *Mathematical Modelling and Computational Experiments*, 1, 407-414.

**Sobol I.M.** (2001) Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. *Mathematics and Computers in Simulation*, 55, 271-280.

59

## Nouvelle méthode

- Une nouvelle méthode basée sur la décomposition d'une fonction a été proposée en 1990 par I. M. Sobol, un mathématicien russe.
- Cette méthode est devenue la méthode de référence mais elle est encore assez peu utilisée:
  - **Nossent J., Elsen P. and Bauwens W.** (2011) Sobol' sensitivity analysis of a complex environmental model. *Environmental Modelling & Software*, 26(12), 1515-1525.
  - **Chevallier, D., Mourrain, B. & Girondot, M.** (2020) Modelling leatherback biphasic indeterminate growth using a modified Gompertz equation. *Ecological Modelling*, 426, 109037



Илья Меерович Соболев

60



## Rappels

- Soit
    - une variable aléatoire  $x$  de variance  $s_x^2$
  - Alors  $s_{k+x}^2 = s_x^2$
  - et  $s_{k \cdot x}^2 = k^2 \cdot s_x^2$
- ```
> x <- rnorm(20, mean=10, sd=2)
> sd(x)^2
[1] 2.12796
> sd(10+x)^2
[1] 2.12796
> sd(10*x)^2
[1] 212.79
```

61

## Partons d'un cas simple

- Supposons que le modèle étudié soit linéaire, et qu'il s'écrive sous la forme suivante :

$$Y = \beta_0 + \sum_{i=1}^p \beta_i X_i.$$

- Si les variables  $X_i$  sont supposées indépendantes, la variance de  $Y$  s'écrit alors :

$$V(Y) = \sum_{i=1}^p \beta_i^2 V(X_i),$$

- où  $\beta_i^2 V(X_i)$  est la part de variance due à la variable  $X_i$ .

62

## L'indice SRC

- La sensibilité de  $Y$  à  $X_i$  peut donc simplement être quantifiée par le rapport de la part de variance due à  $X_i$  sur la variance totale.
- On définit ainsi l'indice de sensibilité SRC (Standardized Regression Coefficient) :

$$SRC_i = \frac{\beta_i^2 V(X_i)}{V(Y)}.$$

63

## L'apport de Ilya Sobol

- Un modèle est décrit comme  $Y=f(X)$  avec  $X$  un vecteur  $\{X_1, X_2, \dots, X_d\}$ .
- Sobol (1990, 1993) a proposé une méthode pour écrire cette fonction  $f(X)$  comme étant une décomposition linéaire :

$$f(\mathbf{X}) = f_0 + \sum_{i=1}^d f_i(X_i) + \sum_{i < j}^d f_{ij}(X_i, X_j) + \dots + f_{1,2,\dots,d}(X_1, X_2, \dots, X_d)$$

avec  $X_i \in [0, 1]$  et  $i=1, 2, \dots, d$

$X$  représente donc un hypercube de côté 1.

64

## $f$ comme une somme de fonctions

- La fonction  $f$  du modèle peut donc être décomposée en une somme de fonctions de dimensions croissantes :

$$\begin{aligned} Y &= f(X_1, \dots, X_p) \\ &= f_0 + \sum_{i=1}^p f_i(X_i) + \sum_{1 \leq i < j \leq p} f_{ij}(X_i, X_j) + \dots + f_{1\dots p}(X_1, \dots, X_p). \end{aligned}$$

où

$$\begin{aligned} f_0 &= \mathbf{E}[Y], \\ f_i(X_i) &= \mathbf{E}[Y|X_i] - \mathbf{E}[Y], \\ f_{i,j}(X_i, X_j) &= \mathbf{E}[Y|X_i, X_j] - \mathbf{E}[Y|X_i] - \mathbf{E}[Y|X_j] + \mathbf{E}[Y], \\ f_{i,j,k}(X_i, X_j, X_k) &= \mathbf{E}[Y|X_i, X_j, X_k] - \mathbf{E}[Y|X_i, X_j] - \mathbf{E}[Y|X_i, X_k] \\ &\quad - \mathbf{E}[Y|X_j, X_k] \dots \end{aligned}$$

65

## Variance du modèle

- La variance du modèle à entrées indépendantes se décompose en :

$$V = \sum_{i=1}^p V_i + \sum_{1 \leq i < j \leq p} V_{ij} + \dots + V_{1\dots p},$$

où

$$\begin{aligned} V_i &= V(\mathbf{E}[Y|X_i]), \\ V_{ij} &= V(\mathbf{E}[Y|X_i, X_j]) - V_i - V_j, \\ V_{ijk} &= V(\mathbf{E}[Y|X_i, X_j, X_k]) - V_{ij} - V_{ik} - V_{jk} - V_i - V_j - V_k, \\ &\dots \\ V_{1\dots p} &= V - \sum_{i=1}^p V_i - \sum_{1 \leq i < j \leq p} V_{ij} - \dots - \sum_{1 \leq i_1 < \dots < i_{p-1} \leq p} V_{i_1 \dots i_{p-1}} \end{aligned}$$

66

## Décomposition linéaire de $f(x)$

- Lorsque la fonction  $f(x)$  est intégrable, il est possible de calculer la décomposition de façon analytique;
- Sinon (ce qui est le plus souvent le cas dans des cas réels), on effectue une décomposition par une méthode itérative basée sur une méthode de (quasi)-Monte-Carlo.

◦ [https://en.wikipedia.org/w/index.php?title=Variance-based\\_sensitivity\\_analysis](https://en.wikipedia.org/w/index.php?title=Variance-based_sensitivity_analysis)

67

## Indices de Sobol

- I. M. Sobol se base sur cette décomposition pour définir des indices de sensibilité d'ordre un:  $S_i = V_i/V$ .
- Les indices de sensibilité d'ordre deux,  $S_{ij} = V_{ij}/V$ , expriment la sensibilité de la variance de  $Y$  à l'interaction des variables  $X_i$  et  $X_j$ , c'est-à-dire la sensibilité de  $Y$  aux variables  $X_i$  et  $X_j$  qui n'est pas prise en compte dans l'effet des variables seules.
- On peut aussi définir des indices d'ordre 3 jusqu'à  $p$ .

68

## Interprétation des indices

- L'interprétation de ces indices est facile, puisque leur somme est égale à 1, et étant tous positifs, plus l'indice sera grand (proche de 1), plus la variable aura d'importance.

69

## Indices composites

- On peut aussi définir la contribution d'une variable à la variance totale comme étant la somme de tous les indices où elle apparaît, par exemple, pour  $S_i$  :

$$S_i + S_{i(1:p)} + S_{i(1:p)(1:p)} + \dots + S_{1:p}$$

*Attention, la somme des  $S_i$  définis de cette façon est supérieure à 1.*

70

## Estimation des indices

- Il existe plusieurs méthodes permettant leur estimation.
- La méthode originale, nommée ici Sobol'93 permet l'estimation de toutes les interactions mais les estimateurs sont assez instables.
- On utilisera aussi celle de Jansen'99 plus stable qui donne la contribution à la variance totale de chaque paramètre pris isolément,  $S_i$ , et de chaque paramètre avec l'ensemble de ses interactions,  $S_{i\cdot}$ .

M.J.W. Jansen, 1999, Analysis of variance designs for model output, Computer Physics Communication, 117, 35–43.

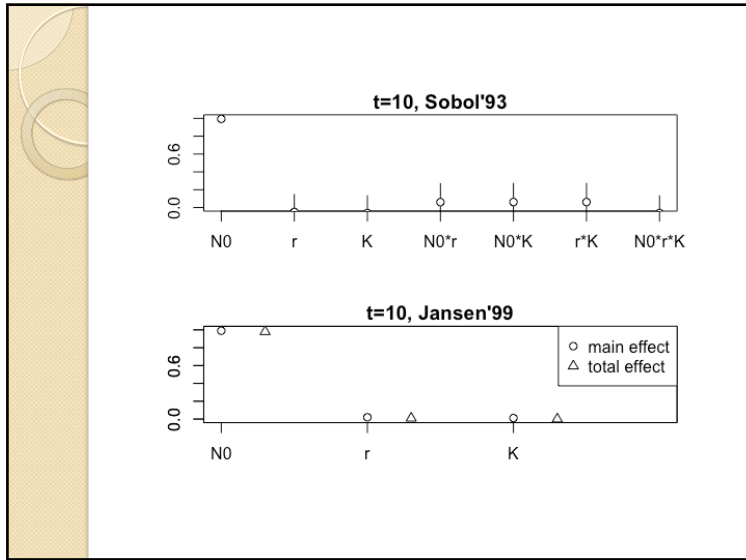
71

## Génération de l'hypercube

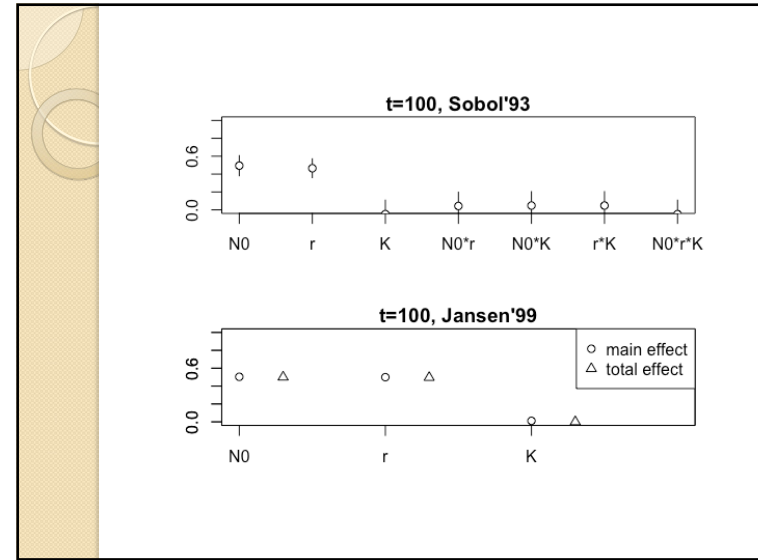
- Les paramètres doivent tous se situer dans un hypercube de côté 1; il convient donc d'effectuer un changement de variable pour les utiliser dans un modèle logistique avec  $N_0=10$ ,  $r=0.01$  et  $K=1000$ :

```
N0 <- ((runif(1)*2)-1)*0.1*10+10
r <- ((runif(1)*2)-1)*0.1*0.01+0.01
K <- ((runif(1)*2)-1)*0.1*1000+1000
```

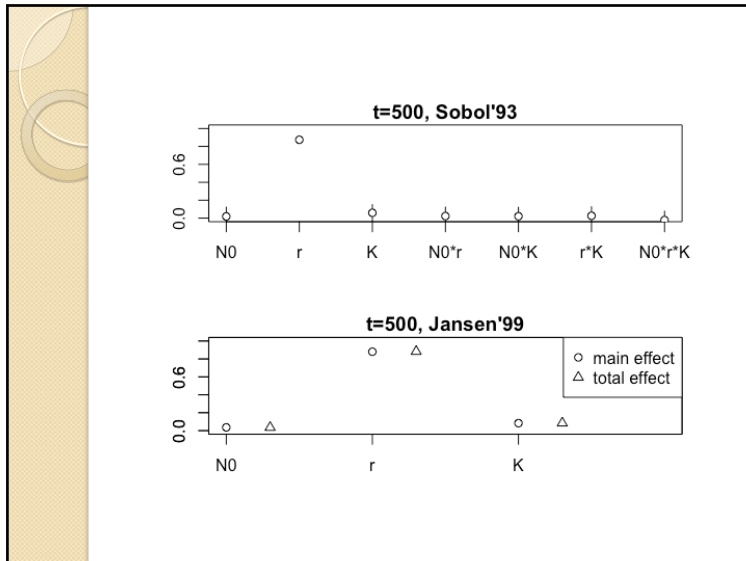
72



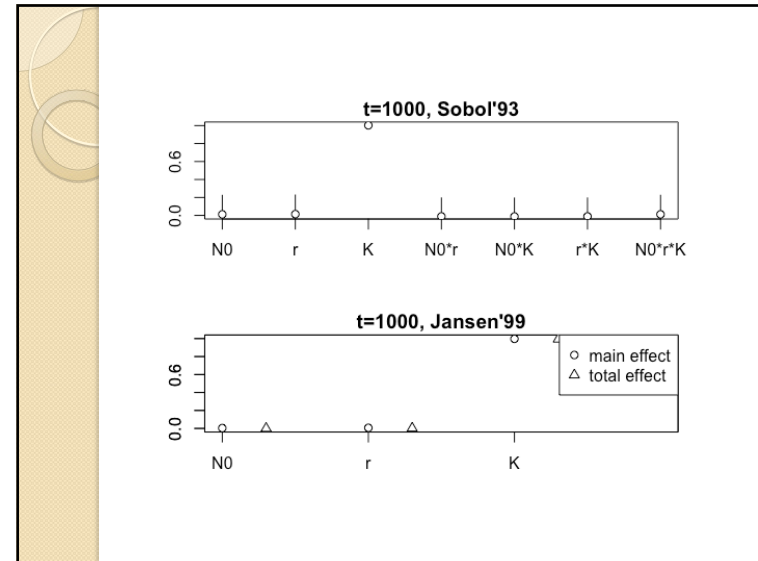
73



74



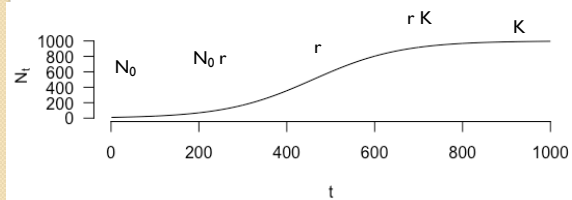
75



76

## Conclusion

- Le modèle est sensible seulement aux effets de premier ordre.
- L'influence des variables diffère selon la zone où on se situe:



77