Revue française d'informatique et de recherche opérationnelle

G. Kreweras

# Suggestibilité individuelle et comportements collectifs

*Revue française d'informatique et de recherche opérationnelle*, tome 2, nº V3 (1968), p. 21-27

<http://www.numdam.org/item?id=RO_1968__2_3_21_0>

R.I.R.O.
(2e année, N° 14, 1968, p. 21-28)

# SUGGESTIBILITE INDIVIDUELLE ET COMPORTEMENTS COLLECTIFS

par G. Kreweras (1)

Résumé. — *L'article reproduit une communication de l'auteur à la réunion d'Amsterdam (European Meeting on Statistics, Econometrics and Management Science, 2-7 sept. 1968), communication qui est elle-même une présentation, dans le langage des choix commerciaux, d'un résultat mathématique antérieurement publié* ]2] *avec référence à une théorie des scrutins à plusieurs tours.*

1. Les considérations développées dans le présent article sont essentiellement méthodologiques. Elles ont pour objet de proposer une définition précise de la notion de *suggestibilité* dans un certain type de contextes, et de déduire de cette définition, avec l'aide d'une propriété mathématique particulière, des conséquences qui pourraient se prêter à des vérifications expérimentales.

Dans le seul but de fixer les idées, nous raisonnons sur une population $\mathcal{C}$ composée de $c$ individus numérotés $i$ ($1 \leqslant i \leqslant c$) ; chacun de ces individus est supposé acheter chaque semaine un journal hebdomadaire et un seul parmi $\omega$ journaux disponibles, l'ensemble des titres de ces journaux étant $\Omega$, et les titres étant numérotés $n$ ($1 \leqslant n \leqslant \omega$).

Nous dirons que l'individu $i$ a un *comportement autonome* soit s'il achète toutes les semaines le même journal, soit plus généralement s'il a une probabilité $q_i(n)$ d'acheter le journal $n$, sans que cette probabilité dépende en quoi que ce soit des journaux achetés par les autres individus.

Pour définir un *comportement hétéronome*, nous pouvons supposer que le kiosque à journaux affiche chaque semaine la liste des nombres d'exemplaires de chaque journal vendus *la semaine précédente* ; si cette liste est

$$(d_1\ d_2 \dots d_n \dots d_\omega) = D,$$

(1) Faculté de Droit et des Sciences Économiques de Paris.

on a évidemment $d_n \geqslant 0$ et

$$d_1 + d_2 + \dots + d_n + \dots + d_\omega = c.$$

Nous nous référons à l'entier $d_n$ comme à la « diffusion du journal $n$ » et à $D$ comme au « vecteur-diffusion » ou plus brièvement à la « diffusion ».

Dans ces conditions nous dirons que l'individu $i$ a un comportement *hétéronome* si sa probabilité d'acheter le journal $n$ est égale à $d_n/c$, rapport qui d'ailleurs ne dépend pas de $i$.

Une autre manière de définir le même type de comportement consiste pour l'individu $i$ à entrer en contact (accidentel et équiprobable) avec n'importe quel individu $j$ de $\mathcal{C}$, à interroger $j$ sur le journal qu'il a acheté la semaine précédente et à imiter aveuglément son choix.

Il est clair que ces deux types de comportement, autonome et hétéronome, sont à certains égards des types extrêmes. Mais nous considérerons en outre des types intermédiaires, qui (par définition) résulteront d'une pondération de ces deux types extrêmes à l'aide de deux coefficients complémentaires $a$ et $b$ $(a + b = 1)$, lesquels seront censés résumer la psychologie de l'individu. Nous appellerons $a_i$ le *degré d'autonomie* et $b_i = 1 - a_i$ le degré d'hétéronomie ou *suggestibilité* de l'individu $i$. De ce fait, la probabilité pour que cet individu $i$ choisisse le journal $n$ deviendra

$$p_i(n) = a_i q_i(n) + b_i \frac{d_n}{c}.$$

Bien entendu si $d_n$ désigne la diffusion relative à la semaine $N$, la probabilité $p_i(n)$ sera relative à la semaine $N + 1$.

2. La connaissance des $c\omega$ nombres $p_i(n)$ permet évidemment, du moins en théorie, de calculer la probabilité pour que le vecteur-diffusion de la semaine $N + 1$ soit un vecteur donné $C$ ; cette probabilité sera, si l'on veut, la probabilité de passage de la diffusion $D$ à la diffusion $C$, probabilité que nous appellerons $m_C^D$.

On définit évidemment ainsi un processus en chaîne de Markov sur l'ensemble $\mathcal{F}$ de tous les « états » possibles, dont chacun est une « diffusion ». $\mathcal{F}$ est l'ensemble de toutes les suites de $\omega$ entiers non-négatifs de somme $c$, et le cardinal de $\mathcal{F}$ est donc, comme on peut le vérifier aisément, le nombre binomial $f = \binom{c + \omega - 1}{\omega - 1}$. De semaine en semaine, la diffusion évolue de l'un à l'autre des $f$ états possibles, qui peuvent être extrêmement nombreux : à titre d'exemple numérique, dans un village où $c = 100$ personnes choisissent leur hebdomadaire parmi $\omega = 4$ titres, on a $f = 176\,851$, ce qui donne une matrice des probabilités de passage de plus de 30 milliards de cases (et chacune des cases nécessiterait en outre un calcul assez long, dont nous n'avons fait ici que constater la possibilité théorique).

Nous avons étudié ailleurs (cf. [1] et surtout [2]) la classe de processus ainsi définie sous le nom de « processus multinomiaux ». Ces processus sont caractérisés par une propriété très remarquable, qui permet d'aboutir à des conclusions intéressantes pour la description qualitative du processus en n'effectuant que des calculs très simples. Nous ne ferons ici qu'énoncer cette propriété en renvoyant à [2] pour sa démonstration complète, et en décrire les conséquences dans le contexte illustratif que nous avons adopté.

3. On sait que dans tout processus de Markov décrit par une matrice des probabilités de passage $M$, la distribution de probabilité sur l'ensemble $\mathcal{F}$ des états au bout de $N$ périodes (ici $N$ semaines) peut être définie par un vecteur-colonne $P_N$ à $f$ composantes non-négatives et de somme 1 ; et que l'on a

$$P_{N+1} = MP_N \qquad (N = 1, 2, ...).$$

On trouve ainsi par multiplication

$$P_N = M^N P_0,$$

le vecteur $P_0$ définissant une « distribution de probabilité initiale » (qui, si l'état initial est connu, a une composante égale à 1 et les autres composantes nulles).

Quoi qu'il en soit ce sont les puissances successives de $M$ qui jouent le rôle central dans l'analyse du processus, et l'étude de ces puissances fait, comme on le sait, intervenir le *spectre* ou ensemble des *valeurs propres* de $M$, c'est-à-dire l'ensemble des racines de l'équation caractéristique $D(\lambda) = 0$, $D(\lambda)$ étant le déterminant de la matrice $M - \lambda I$ (où $I$ désigne la matrice-unité de même ordre $f$ que $M$).

Rappelons que les seules propriétés spectrales tout à fait générales du spectre d'un processus de Markov quelconque sont les suivantes :

1. La somme des ordres de multiplicité des valeurs propres distinctes est égale à $f$.

2. Le spectre est formé de nombres, réels ou complexes, dont les modules sont tous $\leqslant 1$.

3. Le nombre $+1$ appartient au spectre.

Rappelons aussi que parmi les processus de Markov un intérêt particulier s'attache à ceux qui sont *ergodiques* ; ce sont ceux pour lesquels $M^N$ tend vers une matrice limite dont les $f$ colonnes sont identiques à une même colonne $E$. Il est alors facile de voir que le produit $P_N = M^N P_0$ tend précisément vers ce même vecteur-colonne $E$, et cela quel que soit l'état initial $P_0$. Cela signifie que le pronostic probabilisé d'un état futur, éloigné de $N$ périodes d'un état présent connu, tend à devenir, pour des valeurs assez grandes de $N$, à la fois indépendant de $N$ et indépendant de l'état présent $P_0$ ; ce pronostic limite se résume précisément par le *vecteur ergodique* $E$, dont chacune des $f$ composantes donne la probabilité-limite (probabilité ergodique) de l'état correspondant. Une propriété intéressante de ces probabilités ergodiques est que chacune d'elles est la limite vers laquelle tend presque sûrement (c'est-à-dire avec une probabilité égale à 1) la fréquence relative avec laquelle l'état correspondant se rencontrerait si le processus continuait indéfiniment.

Du point de vue des propriétés du spectre, on démontre que le processus est ergodique si et seulement si la valeur propre $+1$ est simple et est la seule valeur propre de module égal à 1. Lorsqu'il en est ainsi le vecteur ergodique $E$ peut se déterminer sans ambiguïté par les deux conditions que $ME = E$ et que les composantes de $E$ aient pour somme 1 ; c'est un *vecteur propre* correspondant à la valeur propre $+1$.

4. La conséquence mathématique essentielle de nos hypothèses consiste en ce que toutes les valeurs propres du processus qu'elles définissent sont assez aisément exprimables à l'aide des paramètres psychologiques $a_i$ et $b_i$ qui caractérisent les acheteurs ; il est commode, en fait, d'utiliser les seuls $b_i$.

Dans ces conditions les nombres $\lambda_0, \lambda_1, \lambda_2, \ldots, \lambda_k, \ldots, \lambda_c$ ci-après sont les valeurs propres, en général distinctes et écrites dans l'ordre décroissant ($\lambda_k > \lambda_{k+1}$) ; en face de chacune d'elles figure l'ordre de multiplicité $f_k$ correspondant (le fait que $f_0 + f_1 + \ldots + f_c = 1$ est une propriété élémentaire des nombres binomiaux) :

$$\lambda_0 = 1 \qquad f_0 = 1$$

$$\lambda_1 = \frac{b_1 + b_2 + \ldots + b_c}{c} \qquad f_1 = \omega - 1$$

$$\lambda_2 = \frac{2!\,(b_1 b_2 + b_1 b_3 + \ldots + b_{c-1} b_c)}{c^2} \qquad f_2 = \binom{\omega}{\omega - 2}$$

$$\lambda_k = \frac{k!\,(b_1 b_2 \ldots b_k + \ldots)}{c^k} \qquad f_k = \binom{k + \omega - 2}{\omega - 2}$$

$$\lambda_c = \frac{c!\,b_1 b_2 \ldots b_c}{c^c} \qquad f_c = \binom{c + \omega - 2}{\omega - 2}$$

La quantité entre parenthèses au numérateur de l'expression de $\lambda_k$ est une somme de $\binom{c}{k}$ termes dont chacun est le produit des suggestibilités de $k$ des $c$ acheteurs.

On remarque en particulier que le spectre est tout à fait indépendant des « comportements spontanés », qui ont été définis par les nombres $q_i(n)$. En outre les valeurs propres elles-mêmes s'expriment exclusivement en fonction des paramètres psychologiques des acheteurs et ne font pas même intervenir le nombre $\omega$ des journaux ; ce dernier intervient par contre dans l'expression des ordres de multiplicité, avec la particularité que s'il n'y a que 2 journaux (et par conséquent $f = c + 1$ états), toutes les valeurs propres sont simples sous la seule condition d'être distinctes.

(Il peut être utile de noter que le cas $\omega = 2$ s'applique peut-être de manière plus réaliste au cas d'un seul journal mais de deux actions possibles : acheter ou ne pas acheter.)

Deux exceptions peuvent se produire relativement à l'inégalité générale $\lambda_k > \lambda_{k+1}$ :

1. Si parmi les $c$ acheteurs il en existe exactement $h$ ($< c$) qui soient parfaitement autonomes, c'est-à-dire pour lesquels $b_i = 0$, alors les $h$ dernières expressions de $\lambda_k$, et elles seules, sont nulles :

$$\lambda_{c-h+1} = \lambda_{c-h+2} = \dots = \lambda_c = 0.$$

0 est alors une valeur propre dont l'ordre de multiplicité est

$$f_{c-h+1} + f_{c-h+2} + \dots + f_c.$$

2. Si tous les $b_i$ sont égaux à 1, c'est-à-dire si tous les acheteurs sont parfaitement hétéronomes, alors non seulement $\lambda_0 = 1$ mais $\lambda_1 = 1$ et par conséquent $+ 1$ est une valeur propre d'ordre de multiplicité

$$1 + (\omega - 1) = \omega.$$

Ce cas, qui est le seul où le processus ne soit pas ergodique, sera considéré séparément.

5. Dans le cas général, où le processus est ergodique, on peut définir de la manière suivante une *diffusion ergodique.* Si $C$ est un état arbitraire du système, c'est-à-dire l'une des $f$ diffusions possibles, on a

$$C = (c_1 \; c_2 \dots c_n \dots c_\omega).$$

Dans cet « état », le journal $n$ est acheté par $c_n$ individus, l'entier $c_n$ pouvant être désigné par $c_n(C)$.

D'autre part le vecteur ergodique $E$ possède une composante $E_C$ relative à ce même état $C$.

On peut alors définir un nombre $\varepsilon_n$ comme la moyenne des $n$-ièmes composantes de tous les états $C$ possibles, avec pondération par les probabilités ergodiques $E_C$ correspondantes :

$$\varepsilon_n = \sum_{C \in \mathcal{F}} E_C c_n(C).$$

Il est facile de voir que les $\omega$ nombres $\varepsilon_n$ ainsi calculés seront non-négatifs et auront pour somme $c$ : ce sont eux qui définiront la diffusion ergodique

$$(\varepsilon_1 \varepsilon_2 \dots \varepsilon_n \dots \varepsilon_\omega).$$

Ces nombres ne définiront pas en général une diffusion proprement dite, c'est-à-dire un état appartenant à l'ensemble $\mathcal{F}$, car ils ne seront en général pas entiers. Toutefois, exactement comme les états, ils donneront lieu au calcul d'un « pourcentage » par journal puisque

$$\frac{\varepsilon_1}{c} + \frac{\varepsilon_2}{c} + \dots + \frac{\varepsilon_n}{c} + \dots + \frac{\varepsilon_\omega}{c} = 1.$$

Ce pourcentage $\varepsilon_n/c$ a du reste une interprétation liée aux propriétés générales des processus ergodiques : il est la limite vers laquelle tend presque sûrement la diffusion moyenne du journal $n$ au cours de $N$ semaines consécutives, lorsque $N$ augmente indéfiniment.

Un fait remarquable est le suivant : le pourcentage $\varepsilon_n/c$ de la diffusion ergodique peut s'obtenir à partir des comportements spontanés $q_i(n)$ par pondération à l'aide des degrés d'autonomie $a_i$ des acheteurs :

$$\varepsilon_n/c = \frac{a_1 q_1(n) + a_2 q_2(n) + \dots + a_c q_c(n)}{a_1 + a_2 + \dots + a_c}.$$

Une hypothèse complémentaire intéressante, dans le cas ergodique, est celle où un certain journal $n$ n'est « représenté » dans l'opinion spontanée d'aucun acheteur ; c'est le cas où $q_i(n)$ est nul pour tout $i$, sauf peut-être pour certains $i$ tels que $a_i = 0$ (en d'autres termes les seuls acheteurs ayant une propension spontanée non nulle à acheter ce journal $n$ sont des individus parfaitement hétéronomes). Il est alors facile de se rendre compte :

1. Qu'il arrivera presque sûrement à ce journal, au bout d'un certain nombre fini de semaines, de n'être acheté par *personne.*

2. Qu'à partir de ce moment ce même journal n'aura plus jamais aucun acheteur, c'est-à-dire se trouvera *définitivement* éliminé.

En particularisant d'une autre manière les hypothèses, on retrouve un cas formellement identique à celui traité complètement dans [1] : c'est le cas où, parmi les $c$ acheteurs, il y en a $r$ qui sont parfaitement autonomes et $s$ qui sont parfaitement hétéronomes ($r + s = \nu$, $r \geqslant 1$). Si l'on suppose en outre que chacun des $r$ individus autonomes achète *toujours* le même journal $n_i$, c'est-à-dire que pour tout $i$ autonome on ait

$$p_i(n) = q_i(n) = \left\{ \begin{array}{l} 1 \; si \; n = n_i \\ 0 \; si \; n \neq n_i \end{array} \right\},$$

alors le résultat général prend la forme suivante : *ergodiquement* les achats des clients hétéronomes se répartissent proportionnellement à la clientèle autonome des différents journaux.

En combinant les hypothèses précédentes, on voit que tout journal $n$ qui ne dispose d'aucun « acheteur inconditionnel » ($r_n = 0$) sera presque sûrement, à la longue, définitivement éliminé, la concurrence se poursuivant exclusivement entre les journaux qui possèdent de tels acheteurs.

Enfin un cas extrême est celui où un seul des journaux, par exemple le journal 1, disposerait de $r_1$ acheteurs inconditionnels, tous les individus autres que ces $r_1$ étant toujours supposés parfaitement hétéronomes.

Il y a, comme précédemment, élimination successive de tous les journaux autres que 1, et par conséquent l'état ergodique est le monopole absolu de ce journal, si petit que soit l'entier $r_1$ pourvu qu'il soit $\geqslant 1$ ; ce cas est à rapprocher du fait que l'obstination d'une très petite minorité peut finir par triompher de l'apathie de la grande masse.

6. Il reste à examiner ce qui se passe si *tous* les individus de $\mathcal{C}$ sont parfaitement hétéronomes, cas dans lequel, on l'a vu, le processus n'est plus ergodique.

Il demeure vrai, dans ce cas, que si un journal n'a par hasard, une certaine semaine, été acheté par personne, ce journal est définitivement éliminé, et que le processus se poursuit entre les journaux restants. On démontre sans trop de difficulté que l'élimination de tous les journaux sauf un se produit presque sûrement à la longue, et qu'il y a donc, comme dans le dernier cas ergodique examiné, tendance vers un monopole final. Mais il est cette fois impossible de prédire avec certitude quel sera le journal bénéficiaire de ce monopole.

Toutefois un pronostic probabilisé est possible sur la base du dernier état connu (état « présent »), et ce pronostic est particulièrement simple : si le dernier état connu est

$$C = (c_1 \ c_2 \ \ldots \ c_n \ \ldots \ c_\omega),$$

avec comme toujours $c_1 + c_2 + \ldots + c_\omega = c$, c'est tout simplement le rapport $c_n/c$ qui se trouve être la probabilité pour que le monopole final se produise au profit du journal $n$.

Cette remarque, ainsi que plusieurs autres aspects du modèle étudié ici, justifie en particulier, vis-à-vis d'une population très hétéronome, les campagnes publicitaires prenant pour thème la diffusion déjà acquise du produit.

Il n'est pas interdit d'espérer que, dans certains cas, on puisse aller un peu plus loin, c'est-à-dire effectuer des mesures au moins grossières de la suggestibilité d'une clientèle, et s'en servir pour choisir les thèmes de publicité et pour prévoir l'ordre de grandeur des délais d'efficacité.

## BIBLIOGRAPHIE

[1] G. Kreweras, *Un modèle d'évolution de l'opinion exprimée par des votes successifs*, Publications de l'Institut de Statistique de l'Université de Paris, vol XII, fasc. 1 (1963).

[2] G. Kreweras, *Spectre des processus multinomiaux et application à certains comportements collectifs*, Publications de l'I.S.U.P., vol. XV, fasc. 2 (1966).