



HAL
open science

Quelques études sur les processus de branchement. Dynamique de populations dans l'asymptotique des grandes déviations

Alain Rouault

► **To cite this version:**

Alain Rouault. Quelques études sur les processus de branchement. Dynamique de populations dans l'asymptotique des grandes déviations. Mathématiques [math]. Université Paris-Sud, 1985. Français. NNT: . tel-04158941

HAL Id: tel-04158941

<https://theses.hal.science/tel-04158941>

Submitted on 11 Jul 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

63788

ORSAY

n° d'ordre :

UNIVERSITE PARIS-SUD

Centre D'Orsay

THESE

De Doctorat D'Etat Es Sciences Mathematiques

présentée pour obtenir le grade de

DOCTEUR ES-SCIENCES

par

Alain ROUAULT

Sujet de la Thèse : Quelques études sur les processus de branchement -
Dynamique de populations dans l'asymptotique des grandes
déviations.

Soutenue le 22 Novembre 1985 devant le Jury composé de :

Didier DACUNHA-CASTELLE	Président
Jean BRETAGNOLLE	
Marie-Claude HEYDEMANN	
Jean-Michel LASRY	
Jacques NEVEU	
Gabriel RUGET.	

Abstract :

First we present two studies in branching processes, one about Zipf's laws, another one about joint estimation of Galton-Watson parameters. Then for spatial branching models, we prove in the large deviation scaling, deterministic limit results for logarithms of populations (generalization of Malthusian parameter, construction of typical genealogies). In the general markovian case, log-populations and log-expectations have not the same limit and an original optimization problem takes place.

Key-words :

Galton-Watson process
Branching random walk
Spatial branching process
Large deviations
Variational formula.
Population dynamics.

A FLORENCE, *bien sûr* ...

Voilà qui est fait ! Je suis arrivé au terme d'un travail long et astreignant. Certains jours, que de découragements, de sollicitations auxquelles je n'échappai pas ! D'autres, au contraire, la joie d'une avancée qui laissait tout espérer !

Tout commença avec l'accueil de Didier Dacunha-Castelle qui m'accorda d'emblée une confiance qu'il sut conserver, me conseilla et me dirigea vers Ruget, pour avoir pressenti mes intérêts mathématiques.

Vint ensuite la clairvoyance exceptionnelle de Gabriel Ruget qui repéra très vite quel intérêt pouvait présenter l'irruption des grandes déviations dans les problèmes de branchement qui m'intéressaient alors. Son enthousiasme me convainquit, nos discussions m'enrichirent et mon travail n'épuise pas aujourd'hui la portée de ses conjectures. Son départ me laissa à la fois orphelin et face à mes responsabilités.

Jean Bretagnolle m'aida. Sa rigueur me fit reprendre une bonne partie de mes démonstrations. Sa compétence lui fit saisir très vite mes difficultés et me permit d'en surmonter plus d'une. Sa disponibilité méritoire m'encouragea à conclure.

Mais il fallut encore pour mener à bien cette recherche d'autres interventions, d'autres rencontres... Monsieur Neveu, dont la présence au jury m'honore, me permit avec gentillesse de profiter de ses résultats sur les arbres. Jean-Michel Lasry me rendit un précieux service en me familiarisant avec les techniques d'optimisation de la manière la plus ouverte et la plus attentive.

Puis Marie-Claude Heydemann me proposa un second sujet avec diligence et efficacité. Les graphes qui jusqu'alors ne m'avaient jamais beaucoup attiré se sont montrés sous un jour nouveau qui n'exclut pas de ma part des excursions ultérieures.

Comment rendre compte de toutes les relations créées au fil des ans au laboratoire de Statistique d'Orsay ? Ma collaboration avec Camille DUBY et Catherine Laredo m'a apporté bien des satisfactions.

.../...

Comment aussi citer tous ceux qui par leur présence amicale, leurs plaisanteries ou leur culture donnèrent à mon travail une dimension moins abstraite et moins solitaire ? Jean-Claude Fort, Patrice Assouad, Dominique Picard, Françoise Garcia parmi tant d'autres, ne manqueront pas de se reconnaître.

La frappe était particulièrement délicate, Mesdames Baillet et Parvan y mirent toute la compétence que nous leur connaissons. La promptitude de Mesdames Launay et Zielinski à l'imprimerie mit un terme à ce travail collectif.

Alain

Selon les lois de l'analyse des sorts, (...) je ne dois point être surpris qu'une chose arrive, lorsqu'elle est possible, et que la difficulté de l'événement est compensée par la quantité des jets.

Diderot (Pensées XXI)

TABLE DES MATIERES

Présentation .	1
A. <u>Estimation non paramétrique de l'espérance et de la variance de la loi de reproduction d'un processus de ramification.</u>	11
Résumé	12
1. Introduction	13
2. Notations	14
3. Approximation de la vraisemblance	15
4. Estimateurs issus de \hat{L} : consistance et normalité conjointe	19
5. Efficacité asymptotique de $\hat{\theta}_{n,z}$	22
B. <u>Quelques rappels et commentaires sur les dualités de Young et Legendre.</u>	27
C. <u>Lois de Zipf et sources markoviennes.</u>	31
Résumé	32
I. Introduction	32
II. Notations	33
III. Lois de Zipf	34
IV. Sources markoviennes	36
1. Définition	36
2. Echantillonnage	37
V. Etude de $Z_n^{m^*}$	41
Supplément à Lois de Sipf et sources markoviennes	52

D. <u>Grandes deviations, dynamique de populations et phénomènes malthusiens.</u>	62
Résumé	63
I. Introduction	64
II. Notations et résultats préliminaires	66
1. Processus de branchement	66
2. Transformée de Cramer	69
3. Encadrements à la Cramer-Chernov	71
III. Résultat principal	73
IV. Discussion	84
1. Deux cas particuliers	84
2. Affaiblissement des hypothèses	86
3. Conjecture sur les âges	88
4. Processus non-homogène	88
E. <u>Lois empiriques dans les processus de branchements spatiaux homogènes supercritiques.</u>	91
I. Position du problème	92
II. Résultats	93
III. Interprétation	95
F. <u>Probabilités de présence dans un processus de branchement spatial markovien.</u>	96
Résumé	97
Introduction	98
§ I. <u>Arbres spatiaux aléatoires</u>	101
Introduction	101
Mesures ponctuelles	101

Espace probabilisable des arbres marqués	102
Construction de l'espace probabilisé	104
Conséquences : mesures de comptage	105
§ II. <u>Noyaux et Cramer-Chernov</u>	107
1. Présentation	107
2. Evaluations sur les tubes	109
§ III. <u>Probabilités de présence - Minoration exactes et lemmes techniques</u>	111
1. Définitions et notations	111
2. Lemme fondamental	112
3. Lemmes techniques	115
§ IV. <u>Probabilités de présence - Résultats asymptotiques</u>	117
1. Hypothèses	117
2. Construction de la famille de processus	118
3. Probabilité de présence - Minoration asymptotique relative à un chemin	118
4. Probabilité de présence - Minoration définitive.	122
G. <u>Espérances et majorations pour un processus de branchement spatial markovien.</u>	128
Résumé	129
Introduction générale	130
I. Etude d'un problème variationnel lié à un modèle de croissance de populations spatiales	133
1. Introduction	133
2. Hypothèses - Existence - Régularité	134
3. Le problème à un pas	139
4. Le problème à n pas	143
5. Résultat annexe	147

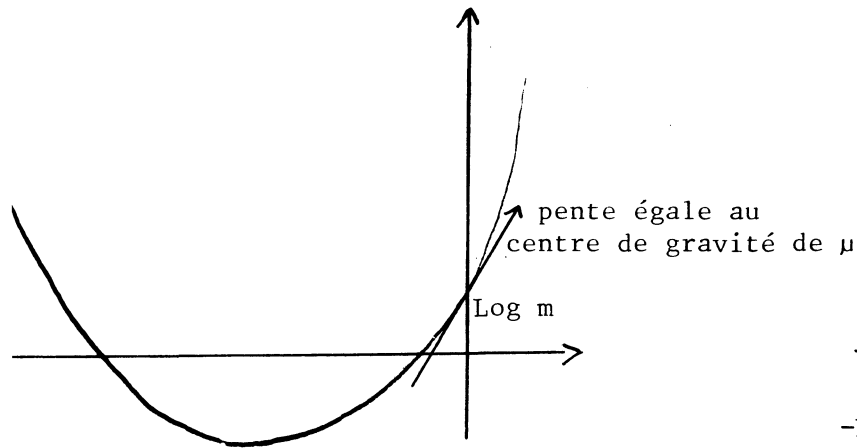
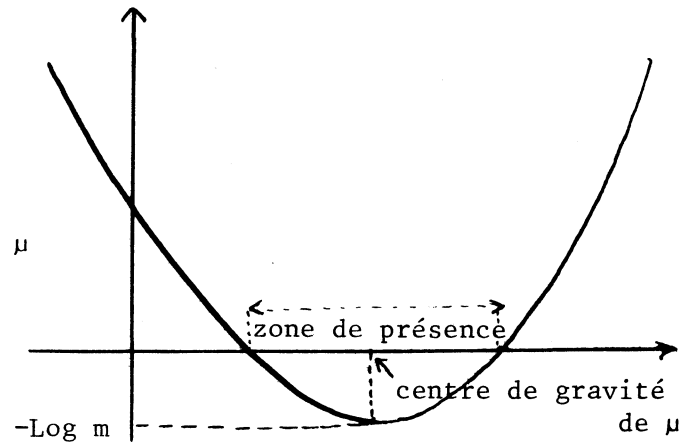
II. Présentation du modèle	149
III. Résultats asymptotiques sur l'espérance du processus	151
1. Minoration	151
2. Majoration	152
IV. Majoration de $\hat{\zeta}_t^3$ en probabilité	158
1. Majoration d'espérances contraintes	158
2. Résultat principal	161
V. Appendice : Preuve de la proposition technique	163
H. <u>Eléments de minoration en zone de présence pour un processus de branchement spatial markovien.</u>	169
Introduction	170
§ 1. Hypothèses et résultats	172
§ II. Outils principaux	175
§ III. Chemin optimal restant sous-critique	178
§ IV. Chemin optimal restant super-critique	180
§ V. Preuve du lemme 2 (équaintégrabilité)	183

PRESENTATION

Le processus de Galton-Watson, un siècle après sa naissance, continue de nourrir une abondante littérature (bibliographies dans [1] et [2]), à laquelle nous avons participé en proposant une estimation non paramétrique simultanée de l'espérance et de la variance [R5]. Par sa nature même, il est le point de passage obligé pour toute modélisation stochastique de croissances exponentielles de populations. L'enrichissement du modèle s'est d'abord développé essentiellement du côté temporel. Deux éléments supplémentaires pouvaient être rendus aléatoires : la durée de vie d'un individu et les dates de naissance relatives de ses enfants. En respectant la règle d'or des processus de branchement - deux individus différents donnent naissance à des processus indépendants "translatés", on a ainsi obtenu les modèles de Bellman-Harris [5], de Sevastyanov [26], de Crump-Mode-Jagers [15]. Le problème principal étant l'effectif Z_t de la population au bout d'un temps t grand, en a fait de façon naturelle une situation relevant de la théorie du renouvellement, et a donné lieu à la définition du paramètre de Malthus ($Z_t \sim e^{\alpha t}$). La question posée (et résolue de manière évidemment astucieuse) par Hammersley en 1974 [14] concerne la répartition dans le temps donc sur \mathbb{R}^+ , des dates de naissance des individus de la n -ième génération, en particulier une limite asymptotique pour la première d'entre elles, normalisée par n . La tentation était grande de passer à \mathbb{R} (ou \mathbb{R}^d) et de considérer le processus temporel précédent comme un processus spatial particulier. A chaque individu x de position b_x on associe une réalisation d'un processus ponctuel sur \mathbb{R} , dont un exemple simple est constitué par le processus de Neymann-Scott [22] : on tire un nombre aléatoire de copies indépendantes d'une même v.a. réelle d'où un ensemble fini de points de \mathbb{R} . On opère une translation de vecteur b_x et on a les positions des enfants de x . On suppose

indépendantes et équidistribuées les réalisations correspondant à différents x . Soit Z_1, \dots, Z_n les mesures aléatoires représentant les générations $1, \dots, n$ et μ la mesure d'intensité du processus (pour B borélien $EZ_1(B) = \mu(B)$). La masse totale m de μ supposée finie et supérieure à 1 représente le nombre moyen d'enfants par individu et $\mu_0 = m^{-1}\mu$ est une probabilité. L'égalité $EZ_n(B) = \mu^{*n}(B) = m^n \mu_0^{*n}(B)$ met bien en évidence la nature des 2 phénomènes stochastiques sous-jacents : Galton-Watson et marche aléatoire. La normalisation de l'espace par n , introduite par la problématique de la première naissance, rend nécessaire l'emploi des grandes déviations et du même coup atténue un peu la portée de la plupart des études spatiales antérieures dont les normalisations relevaient du théorème de la limite centrale. En effet si a est plus grand que le centre de gravité de μ (espérance de μ_0), $EZ_n [n(a, a+\delta)]$ est le résultat de la compensation entre la croissance exponentielle de la population (m^n) et la décroissance de la probabilité ($e^{-n h_0(a)}$ si h_0 est la transformée de Cramer [3] de μ_0), ce qui permet de définir une zone de présence asymptotique. C'est à Biggins [7]...[13], avec l'impulsion de Kingman [16] que revient le mérite d'avoir caractérisé cette zone et décrit ce qui s'y passe. Nous énonçons ici le résultat le plus simple. On suppose la transformée de Laplace $\hat{\mu}$ de μ définie par $\hat{\mu}(\theta) = \int e^{\theta x} \mu(dx)$, finie dans un voisinage de 0, et on pose $L = \text{Log } \hat{\mu}$ et $h = \sup_{\theta} \theta x - L(\theta)$ la duale de Young de L (qui est convexe). Pour tout (a, δ) de $\mathbb{R} \times \mathbb{R}_*^+$, si Δ_a^δ désigne $[a, a+\delta)$ ou $(a-\delta, a]$ suivant le signe de $h'(a)$ (h est finie et analytique dans l'intérieur de l'enveloppe convexe du support de μ) on a

$$\begin{aligned} & \lim_{n \rightarrow \infty} \frac{1}{n} \text{Log } E Z_n (n \Delta_a^\delta) = -h(a) \\ \text{si } h(a) < 0 & \quad \lim_{n \rightarrow \infty} \frac{1}{n} \text{Log } Z_n (n \Delta_a^\delta) = -h(a) \quad \text{p.s. sur } \{Z_n(\mathbb{R}) \rightarrow +\infty\} \\ \text{si } h(a) > 0 & \quad \lim_{n \rightarrow \infty} Z_n (n \Delta_a^\delta) = 0 \quad \text{p.s.} \end{aligned}$$

Graphe de L Graphe de $h = h_0 + \text{Log } m$ 

La fonction $-h$, en zone de présence, joue le rôle de $\text{Log } m$ pour le processus de Galton-Watson supercritique.

Il était dès lors naturel d'envisager un certain nombre de généralisations, tant du côté branchement (multitype, Bellman-Harris, autodépendant) que du côté marche aléatoire (Markov). A la lumière de Ventsel [27] et Azencott-Ruget [4] le cadre adéquat pour asseoir les futurs résultats consiste à se donner un paramètre d'échelle, ε , destiné à tendre vers 0 et fixant les instants de reproduction et la taille des sauts. Une relecture de Biggins en ces termes donne près de x à l'instant t une population en $\exp \frac{1}{\varepsilon} \{-th(\frac{x}{t}) + o(1)\}$ si $h(\frac{x}{t}) < 0$ et 0 sinon. La fonction $\alpha(t, x) = -th(\frac{x}{t})$ vérifie l'équation aux dérivées partielles (e.d.p.) non linéaire $\frac{\partial \alpha}{\partial t} = L(-\frac{\partial \alpha}{\partial x})$, classique dans les problèmes de lois de conservation [19]. L'objectif devient désormais, dans les modèles de branchement spatiaux, la recherche d'" ε -log densités" et l'étude éventuelle des e.d.p. associées. Le modèle de Crump-Mode-Jagers, auquel on rajoute une dimension spatiale, supporte très bien le traitement [R6] mettant en évidence une fonction homogène α de t et de x , l'e.d.p. associée $L(-\frac{\partial \alpha}{\partial t}, -\frac{\partial \alpha}{\partial x}) = 0$ généralisant l'équation donnant le paramètre de Malthus. Cette démonstration donne, au passage une interprétation de toute situation mêlant croissance exponentielle et renouvellement. Elle a fourni, a posteriori, l'idée d'une démonstration alternative et éclairante de résultats antérieurs [R1] sur une modélisation probabiliste des lois de Zipf, d'origine linguistique.

La conjecture la plus générale de Ruget [24][25] porte sur un processus de branchement spatial dont la loi de reproduction dépend du temps, de l'espace et de la population elle-même, c'est-à-dire notamment que le branchement sous-jacent est auto-dépendant (ou contrôlable ou dépendant de la taille) et la marche sous-jacente markovienne. Mon propos était d'avancer dans cette direction.

Dans un premier temps on se donne une famille indexée par $\mathbb{R}^+ \times \mathbb{R}$ (temps, espace) de lois de probabilité sur M , espace des mesures ponctuelles finies sur \mathbb{R} . L'élément générique est $P_{t,x}(dv)$, ayant une espérance notée $\mu_{t,x}$, mesure de Radon positive sur \mathbb{R} ayant une log-Laplace $L(t,x,\cdot)$ et une transformée de Cramer $h(t,x,\cdot)$. $P_{t,x}$ décrit, à l'échelle 1, la loi des positions des enfants d'un individu situé en x à l'instant t , repérées à partir de x . Si besoin est, on particularisera $P_{t,x}$ en supposant qu'il est de Neymann-Scott. Le cas homogène correspond à $P_{t,x} \equiv P$ pour tout t,x . Pour tout $\varepsilon > 0$ on se donne π_0^ε élément de M . Chaque élément x_0 du support de π_0^ε doit être pensé comme la position d'un ancêtre (individu de la génération 0). Il vit un temps ε puis est remplacé par des enfants dont les positions forment un élément (aléatoire) de M obtenu de la manière suivante : on tire un élément de M suivant P_{ε,x_0} , on lui applique une homothétie de rapport ε puis une translation de vecteur x_0 . Les enfants de l'ancêtre forment la première génération et sont datés de l'instant ε , de sorte qu'un tel individu, situé en x , sera à l'instant 2ε remplacé par des enfants dont les positions formeront un élément de M obtenu par tirage suivant $P_{2\varepsilon,x}$, ε -homothétie et x -translation. On itère la procédure et on s'intéresse pour tout t à l'élément ζ_t^ε de M formé par les individus en vie à t (donc de génération $[\frac{t}{\varepsilon}]$).

On définit pour tout ε , la famille mesurable $Q_t^\varepsilon(x,\cdot)$, $(t,x) \in \mathbb{R}_+ \times \mathbb{R}$ de probabilités sur M par

$$Q_t^\varepsilon(x,dv) = T_x S_\varepsilon P_{t,x}(dv)$$

où $S_\varepsilon f(z) = f(\varepsilon z)$ et $T_x f(z) = f(z+x)$ - opérateurs d'échelle et de translation -,

puis on construit la chaîne de Markov ζ_n , $n \in \mathbf{N}$ d'espace d'états M , d'état initial π_0^ε , avec la propriété de branchement : pour n fixé, conditionnellement en ζ_0, \dots, ζ_n la loi de ζ_{n+1} est celle de la superposition de k v.a. indépendantes de lois respectives $Q_{(n+1)\varepsilon}^\varepsilon(x_j, \cdot)$ si $\zeta_n = \sum_{j=1}^k \delta_{x_j}$ (où δ_x désigne la mesure de Dirac en x). Pour toute f continue à support compact on a

$$\iint_M \nu(f) Q_{t\varepsilon}^\varepsilon(x, d\nu) = T_x S_\varepsilon \mu_{t,x}^\varepsilon(f) \text{ noté } Q_t^\varepsilon f(x)$$

et pour les tribus propres,

$$E^\varepsilon \zeta_{n+1}(f) = \zeta_n(Q_{(n+1)\varepsilon}^\varepsilon f)$$

d'où on tire si pour $t \in \varepsilon \mathbf{N}$, μ_t^ε désigne l'espérance de ζ_t^ε

$$\mu_t^\varepsilon = \mu_{t-\varepsilon}^\varepsilon Q_t^\varepsilon$$

La structure de Q_t^ε et la propriété précédente met en évidence l'analogie avec un problème de grandes déviations pour chaînes de Markov [4] [27] (un seul individu par génération : $\zeta_t^\varepsilon = \delta_{U_t^\varepsilon}$ et les résultats portent sur $\varepsilon \log P(U_t^\varepsilon \in I)$). Cette étude d'espérance avec les outils précédents et tenant compte d'une population initiale en $\exp \frac{f_0(x_0)}{\varepsilon}$ près de x_0 tel que $f_0(x_0) > 0$, fournit près de x à t la quantité $\exp \frac{\beta(t,x)}{\varepsilon}$ où

$$\beta(t,x) = \sup_{\varphi} f_0(\varphi(0)) - \int_0^t h(s, \varphi(s), \dot{\varphi}(s)) ds$$

$\varphi(t)=x$

On retrouve la problématique des tubes autour de chemins minimisant l'intégrale d'action. Un chemin doit être vu comme la limite (en ε) d'une ligne généalogique constituée par les positions des ancêtres successifs d'un individu. Dans le cas homogène, vu la convexité de h , les chemins optimaux sont rectilignes.

Un certain nombre d'étapes sont apparues, certaines contingentes, d'autres nécessaires, nous obligeant à revenir sur le cas homogène pour comprendre certaines situations de base. Dans la première catégorie se situe notre étude [R3] de la loi empirique du processus de reproduction, calculée le long d'une ligne généalogique

et conditionnée par l'arrivée près d'un lieu donné (différent du centre de gravité). En ce qui concerne la seconde, on remarque que l'existence d'une population initiale exponentiellement grande est compatible avec un processus qui, démarrant avec un seul individu, serait sous-critique ; ceci nécessite donc l'étude des probabilités de présence [R7], [R8] , qui même dans le cas homogène fait appel à des idées issues des outils markoviens. En chemin la nécessité de construire correctement l'espace de probabilité pour un branchement spatial markovien nous a fait rencontrer les arbres marqués de Neveu [21] .

Il restait à attaquer la comparaison entre ζ_t^ε et son espérance. Dans le cas homogène, la linéarité des chemins optimaux - par ailleurs caractéristiques de l'équation aux dérivées partielles - a pour conséquence qu'un tel chemin aboutissant en x appartenant à la zone de présence à l'instant t n'a jamais avant t quitté la zone de présence correspondante. C'est ici qu'apparaît la spécificité du cas non-homogène. L'espérance et la variable aléatoire ne sont plus équivalentes en général, la conjecture [24] , [25] étant que la v.a. correspond à une formule variationnelle à contrainte permanente, rendant compte de l'extinction des lignées à partir de la sortie de la zone de présence. Plus précisément

$\varepsilon \text{Log } \zeta_t^\varepsilon ([x-r(\varepsilon), x+r(\varepsilon)])$ tend pour un $r(\varepsilon)$ convenable vers $\alpha(t,x) = \sup_{\varphi} T(\varphi,t)$ où $T(\varphi,t) = f_0(\varphi(0)) - \int_0^t h(s,\varphi(s),\dot{\varphi}(s)) ds$ et le sup est pris parmi les φ tels que $\varphi(t) = x$ et pour tout $0 \leq s \leq t$, $T(\varphi,s) > 0$.

Notre contribution à la preuve de cette conjecture a consisté en une étude séquentielle. Il est utile de découper $[0,t]$ en tranches d'épaisseur τ suffisamment faible pour que qualitativement tout se passe comme pour un processus homogène (v.a. \sim esp. lorsque celle-ci est > 1) mais suffisamment grande pour que le processus ait eu le temps de donner des résultats asymptotiques ($\varepsilon \ll \tau \ll 1$) . Il est bien entendu que les points de départ spatiaux d'une tranche sont les points d'arrivée de la tranche précédente ayant une densité de population suffisante. L'étude d'une tranche nécessite notamment un encadrement de l'espérance correspondante

donc les mêmes outils que dans l'espérance globale. Mais l'empilement d'un nombre croissant ($\sim \frac{t}{\tau}$ à t fixé) de ces tranches exige de ne pas s'en tenir aux parties principales quand $\varepsilon \rightarrow 0$, c'est-à-dire d'évaluer les restes de façon uniforme. Cette étude a donc été faite, dans cet esprit, une fois pour toutes [R9] .

L'évaluation de restes dans les inégalités a requis des extrémales à pentes bornées, et même bornées uniformément . L'interférence permanente entre analyse convexe, équation de Hamilton-Jacobi [20] (β vérifie $\beta(0, \cdot) = f_0$, $\frac{\partial \beta}{\partial t} = L(t, x, -\frac{\partial \beta}{\partial x}$) et en est la solution "entropique") et grandes déviations markoviennes trouve sa trace dans les hypothèses présentées [R8 , R9] , à la fois proche de Benton [6] et de Ventse1 [27].

Il reste dans ce modèle (dit) markovien à parachever la minoration du chap.H et montrer la convergence du modèle τ -approché (et rendre compte de certaines irrégularités de la correspondance entre f_0 et α [18]). Les résultats sur Hamilton-Jacobi avec L dépendant aussi de β ouvrent des perspectives de démonstration dans le cas auto-dépendant qui décrit assez bien la croissance de populations réelles souvent tempérée par des saturations, pour lequel la pertinence de la conjecture [24] a déjà été mise en évidence par les simulations de Laredo [17] et Pierre-Loti-Viaud [23] . Il faut cependant noter que dans le modèle général, L dépendra de α et non de β (de la population et non de son espérance).

Références :

- [1] ASMUSSEN (S.) - HERING (H.), *Branching processes*. Birkhäuser, London, 1983.
- [2] ATHREYA (K.B.) - NEY (P.E.), *Branching processes*, Springer, Berlin, 1972.
- [3] AZENCOTT (R.), *Cours de l'Ecole d'Eté de Saint-Flour, Lecture Notes in Math. n° 774 (1978)*.
- [4] AZENCOTT (R.) - RUGET (G.), *Mélanges d'équations différentielles et grands écarts à la loi des grand nombres. Zeits. für Wahr., 1977, t.38, p.1-54*.
- [5] BELLMAN (R.) - HARRIS (T.E.), *On age-dependent binary branching processes. Ann. of Math. 1952, t.55, p. 280-295*.
- [6] BENTON (S.), *The Hamilton-Jacobi Equation : A global approach. Academic Press, New York, 1977*.
- [7] BIGGINS (J.D.), *The first- and last - birth problems for a multitype age-dependent branching process. Adv. Appl. Prob. 1976, t.8, p.446-459*
- [8] BIGGINS (J.D.), *Martingale convergence in the branching random walk. Journ. of Appl. Prob. 1977, t. 14, p. 25-37*.
- [9] BIGGINS (J.D.), *Chernoff's theorem in the branching random walk. Journ. of Appl. Prob. 1977. t.14, p. 630-636*.
- [10] BIGGINS (J.D.), *The asymptotic shape of the branching random walk. Adv. Appl. Prob. 1978, t.10, p. 62*.
- [11] BIGGINS (J.D.), *Growth rates in the branching random walk. Zeits. für Wahrs. 1979, t. 48, p. 17-34*.
- [12] BIGGINS (J.D.), *Spatial Spread in branching processes, in "Biological growth and spread". Lecture Notes in Biom. 1979, t.38*.
- [13] BIGGINS (J.D.), *Limiting point processes in the branching random walk. Zeits für Wahrs. 1981, t.55, p. 297-303*.

- [14] HAMMERSLEY (J.M.), *Postulates for subadditive processes*. *Ann. Prob.* 1974, t.2, p. 652-680.
- [15] JAGERS (P.), *Branching processes with biological applications*. Wiley, London, 1975.
- [16] KINGMAN (J.F.C.), *The first birth problem for an age-dependent branching process*. *Ann. Prob.* 1975, t.12, p. 341-345.
- [17] LAREDO (C.), *Quelques études asymptotiques en dynamique de populations 1981, Thèse de 3ème cycle, Paris-Sud*.
- [18] LAREDO (C.), MOREL (J.M.), *Propriétés de stabilité de l'opérateur décrivant la répartition spatio-temporelle ...*, *Comptes-Rendus A.S.* 1984.
- [19] LAX (P.D.), *Hyperbolic systems of conservation laws, II*, *Comm. Pure Appl. Math.* 1957, t.10, p. 537-566.
- [20] LIONS (P.L.), *Generalized solutions of Hamilton-Jacobi equations*. Pitman, London, 1982.
- [21] NEVEU (J.), *Arbres et processus de Galton-Watson (à paraître)*.
- [22] NEYMANN (J.) - SCOTT (E.L.), *Processes of clustering and applications, in Stochastic Point Processes edit. P.A.W. Lewis, Wiley, New York, 1972*.
- [23] PIERRE-LOTI-VIAUD (D.), *Etude de phénomènes hystérétiques en dynamique de populations, 1984, Thèse de 3ème cycle, Paris-Sud*.
- [24] RUGET (G.), *Large deviations and more or less rare events in population dynamics. Colloque Luminy. Lect. Notes in Biom.* 1981, t.49, p. 388.
- [25] RUGET (G.), *Grandes déviations et dynamique des populations. Eleventh Conference on "Stoch. Proc. and Their Appl." in (idem), 1983, t.16, p. 10-11*.
- [26] SEVASTYANOV (B.A.), *Branching processes*. Nauka. Moscou, 1971.
- [27] VENTSEL (A.D.), *Rough limit theorems on large deviations for Markov stochastic processes. Theor. of Prob. and its Appl.* 1976, t.21, n°2, p. 227-242 et t.21, n° 3, p. 499-512.

Références de l'auteur :

- [R1] *Lois de Zipf et sources markoviennes*
Annales de l'IHP. XIV n° 2, 169-188 (1978) et XV n° 2, 195-196 (1979)
- [R2]* *Estimation simultanée de l'espérance et de la variance pour un processus de Galton-Watson (en collaboration avec C. Duby)*
Note aux CRAS, 290, 339-341 (1980)
- [R3] *Lois empiriques dans les processus de branchement spatiaux*
Note aux CRAS, 292, 933-936 (1981)
- [R4]** *Une généralisation du paramètre de Malthus pour les processus de Crump-Mode-Jagers spatiaux (en collaboration avec C. Laredo)*
Note aux CRAS, 193, 481-484 (1981)
- [R5] *Estimation simultanée de l'espérance et de la variance pour un processus de Galton-Watson (en collaboration avec C. Duby)*
Annales de l'IHP, XVIII, n° 2, 149-163 (1982)
- [R6] *Grandes déviations, dynamique de populations et phénomènes malthusiens (en collaboration avec C. Laredo)*
Annales de l'IHP, XIV, n° 4, 323-350 (1983)
- [R7] *Comportement asymptotique d'un processus de branchement spatial en zone sous-critique*
Note aux CRAS, 299, 887-890 (1984)
- [R8] *Probabilités de présence dans un processus de branchement spatial markovien (soumis aux CRAS et aux Annales de l'IHP)*
- [R9] *Espérances et majorations pour les processus de branchement spatiaux markoviens (soumis aux Annales de l'IHP)*

* Annonce de [5]

** Annonce de [6]

A. Estimation non paramétrique de l'espérance et de la variance de la loi de reproduction d'un processus de ramification.

Estimation non paramétrique de l'espérance et de la variance de la loi de reproduction d'un processus de ramification

par

Camille DUBY

Laboratoire de Biométrie, I. N. R. A., 78000 Versailles

et

Alain ROUAULT

Mathématiques, Bât. 425, E. R. A.-C. N. R. S. 532,
Université de Paris-Sud, 91405 Orsay Cedex

RÉSUMÉ. — On étudie l'estimation conjointe de la moyenne m et de la variance σ^2 de la loi de reproduction d'un processus de Galton-Watson dans le cas $m > 1$, $p_0 = 0$, quand les observations sont les tailles successives des générations, $Z_0 = z, Z_1, \dots, Z_n$. Quand n et z tendent vers l'infini, on approche la log-vraisemblance. Le problème devient paramétrique. On montre la normalité asymptotique et l'optimalité du couple $(\hat{m}, \hat{\sigma}^2)$ ainsi déduit.

SUMMARY. — The estimations of the mean m and variance σ^2 of the offspring distribution of a Galton-Watson process are studied jointly in the case $m > 1$, $p_0 = 0$. Observations are sizes of generations, $Z_0 = z, Z_1, \dots, Z_n$. For z and $n \rightarrow \infty$ the log-likelihood is approximated. The problem becomes a parametric one and it is shown that the pair $(\hat{m}, \hat{\sigma}^2)$ is asymptotically normal and optimal.

1. INTRODUCTION

On étudie, pour les processus de Galton-Watson, l'estimation de l'espérance m et de la variance σ^2 du nombre Y de descendants d'un individu dans le cas où $m = EY > 1$, $p_0 = P(Y = 0) = 0$ et où l'échantillon est constitué des tailles des générations successives $Z_0 = z, Z_1, \dots, Z_n$.

En présentant les résultats connus sur ces estimations, Dion et Keiding [5] examinent deux passages à la limite : en n et/ou en z . L'estimateur de Harris

$$\hat{m} = \sum_{k=1}^n Z_k / \sum_{k=1}^n Z_{k-1}$$
 maximise la vraisemblance dans le modèle complé-

ment général, indexé par $\{p = (p_k)_{k \in \mathbb{N}^*}, \sum p_k = 1\}$, ou encore dans le modèle « conditionnellement exponentiel » à un paramètre. A z fixé et $n \rightarrow \infty$, il est consistant et asymptotiquement normal, mais avec une normalisation aléatoire. Dans le modèle paramétrique cité, il est optimal (Heyde [9]). Si z et $n \rightarrow \infty$, \hat{m} est consistant et asymptotiquement normal, avec normalisation déterministe. Il existe des modèles paramétriques pour lesquels \hat{m} n'est pas l'estimateur du maximum de vraisemblance.

Vu la propriété $Z_{n+1} \stackrel{\mathcal{L}}{=} \sum_1^{Z_n} \xi_{j,n}$ (où les $\xi_{j,n}$ sont des v. a. indépendantes, de même loi, la loi de reproduction), un estimateur naturel de la variance

$$\hat{\sigma}^2 = n^{-1} \sum_{k=1}^n (Z_k - \hat{m}Z_{k-1})^2 (Z_{k-1})^{-1}.$$
 On sait qu'il est

convergent et asymptotiquement normal à z fixé quand $n \rightarrow \infty$.

On a cherché ici à aborder l'estimation de m et σ^2 conjointement, sans modèle paramétrique souvent difficile à justifier dans les applications.

On part de l'écriture

$$(1.1) \quad Z_{n+1} = mZ_n + \sigma\sqrt{Z_n}\varepsilon_{n+1}$$

La normalité asymptotique des ε_n (Heyde-Brown [11]) suggère d'approcher la log-vraisemblance L de l'échantillon par

$$(1.2) \quad \hat{L} = -\frac{n}{2} \log(2\pi\sigma^2) - \frac{1}{2} \sum_{k=1}^n \log Z_{k-1} - \frac{1}{2} \sum_{k=1}^n (Z_k - mZ_{k-1})^2 \sigma^{-2} (Z_{k-1})^{-1}$$

m et σ^2 sont les seuls paramètres qui subsistent et il se trouve que \hat{L} est maximisé par \hat{m} et $\hat{\sigma}^2$. Ces deux estimateurs ont des vitesses de convergence très différentes (l'une en $\sqrt{m^nz}$, l'autre en \sqrt{n}) ce qui rend difficile une étude conjointe. C'est à l'aide de \hat{L} qu'on a pu étudier leurs propriétés asymptotiques, normalité et efficacité. Ceci exige une évaluation assez fine de $L - \hat{L}$. Si seul $n \rightarrow \infty$ l'approximation est trop grossière pour pouvoir déduire les propriétés habituelles des estimateurs du maximum de vraisemblance. Il est nécessaire d'avoir à la fois n et $z \rightarrow \infty$, ce qui signifie que l'information utile du processus est contenue dans ses générations éloignées, information s'exprimant à travers les seuls paramètres m et σ^2 . Le cas où seul $z \rightarrow \infty$ donne une bonne approximation mais $\hat{\sigma}^2$ ne converge plus.

Après avoir défini les notations dans la 2^e partie, on étudie dans la 3^e partie l'approximation de la vraisemblance \hat{L} et l'évaluation du reste $L - \hat{L}$. Dans la 4^e partie, à l'aide de \hat{L} on montre la normalité (et l'indépendance) asymptotique du couple $(\hat{m}, \hat{\sigma}^2)$ et dans la 5^e partie son efficacité asymptotique au sens de Weiss et Wolfowitz.

Ces résultats ont été annoncés dans [6].

2. NOTATIONS

Soit $a = (a_k)_{k \in \mathbb{N}^*}$, $b = (b_k)_{k \in \mathbb{N}^*}$ deux suites de nombres positifs non tous nuls telles que

$$(2.1) \quad \forall k \ a_k \leq b_k \quad \sum_1^{\infty} k^4 b_k < \infty \quad \sum_1^{\infty} k a_k > 1$$

$$(2.2) \quad \text{PGCD} \{ k \mid a_k \neq 0 \} = 1$$

$$(2.3) \quad \mathcal{P} = \left\{ p = (p_k)_{k \in \mathbb{N}^*} \mid \forall k \ a_k \leq p_k \leq b_k, \sum_1^{\infty} p_k = 1 \right\} \neq \emptyset$$

Notons pour $p \in \mathcal{P}$:

$$(2.4) \quad m = \sum_1^{\infty} k p_k, \quad \sigma^2 = \sum_1^{\infty} (k - m)^2 p_k$$

$$\theta = (m, \sigma^2), \quad \mu^3 = \sum_1^{\infty} |k - m|^3 p_k$$

$$\Omega = \prod_{i,j=1}^{\tau} \mathbb{N}^*, \quad \mathcal{F} = \bigotimes_{i,j=1}^{\tau} \mathcal{P}(\mathbb{N}^*), \quad P = \bigotimes_{i,j=1}^{\tau} p$$

$Y(i, j)$ la (i, j) ^{ième} coordonnée et pour tout i

$$(2.5) \quad \begin{aligned} Z_0^i &= 1, & Z_1^i &= Y(i, 1), & Z_2^i &= Y(i, 2) + \dots + Y(i, 1 + Z_1^i), \dots \\ Z_n^i &= Y(i, 2 + Z_1^i + \dots + Z_{n-2}^i) + \dots + Y(i, 1 + Z_1^i + \dots + Z_{n-1}^i) \\ Z_{k,z} &= \sum_{i=1}^z Z_k^i, & \mathcal{F}_n^i &= \sigma(Z_k^i, 0 \leq k \leq n), & \mathcal{F}_{n,z} &= \sigma(\mathcal{F}_n^i, 1 \leq i \leq z). \end{aligned}$$

On pose aussi :

$$(2.6) \quad \varepsilon_{n,z} = (Z_{n,z} - mZ_{n-1,z})\sigma^{-1}(Z_{n-1,z})^{-1}$$

Dans ces conditions, le processus $(Z_{n,z})_{n \in \mathbb{N}}$ est un processus de Galton-Watson avec z ancêtres de loi de reproduction p , supercritique, avec explosion presque sûre. De plus pour $p \in \mathcal{P}$, m, σ^2, μ_3 sont uniformément majorés et minorés.

3. APPROXIMATION DE LA VRAISEMBLANCE

Le processus étant markovien, la log-vraisemblance de l'échantillon est

$$(3.1) \quad L_{n,z} = \sum_{k=1}^n \log P(Z_{k,z} | Z_{k-1,z})$$

En utilisant la loi limite des ε (cf. [11]), on l'approche par

$$(3.2) \quad \hat{L}_{n,z} = -\frac{n}{2} \log(2\pi\sigma^2) - \frac{1}{2} \sum_{k=1}^n \log Z_{k-1,z} - \frac{1}{2} \sum_{k=1}^n (\varepsilon_{k,z})^2$$

et on pose :

$$(3.3) \quad h_{n,z} = L_{n,z} - \hat{L}_{n,z}$$

PROPOSITION 1. — Pour toute g telle que $g(z) \rightarrow +\infty$ quand $z \rightarrow \infty$, $h_{n,z}$ est un $o_P(g(z)z^{-\frac{1}{2}})$ quand $n, z \rightarrow \infty$ et ceci uniformément dans \mathcal{P} .

On a

$$(3.4) \quad h_{n,z} = \sum_{k=1}^n \log \left[1 + \frac{1}{\sqrt{2\pi}} \Delta_{k,z} \exp \frac{1}{2} (\varepsilon_{k,z})^2 \right]$$

où

$$(3.5) \quad \Delta_{k,z} = 2\pi\sigma(Z_{k-1,z})^{\frac{1}{2}} \left[P(Z_{k,z} | Z_{k-1,z}) - \sigma^{-1}(2\pi Z_{k-1,z})^{-\frac{1}{2}} \exp -\frac{1}{2}(\varepsilon_{k,z})^2 \right]$$

Pour évaluer $h_{n,z}$ on contrôle $\Delta_{k,z}$ en adaptant un théorème limite local (Petrov [15]) pour obtenir une majoration uniforme dans \mathcal{P} .

PROPOSITION 2. — Il existe A positif, tel que pour tout p de \mathcal{P} et a, b, k, z de \mathbb{N}^*

$$(3.6) \quad \left| P(Z_{k,z} = b | Z_{k-1,z} = a) - \frac{1}{\sigma\sqrt{2\pi a}} \exp\left(-\frac{1}{2} \frac{(b - ma)^2}{a\sigma^2}\right) \right| \leq \frac{A}{a}$$

Admettons-la pour l'instant.

Démonstration de la proposition 1. — Compte tenu de (3.6), il suffit de montrer le même résultat sur

$$(3.7) \quad \tilde{h}_{n,z} = \sum_{k=1}^n (Z_{k-1,z})^{-\frac{1}{2}} \exp \frac{1}{2} (\varepsilon_{k,z})^2$$

On utilise alors une méthode du genre « logarithme itéré ». Pour $\delta > 0$ fixé on pose

$$(3.8) \quad A_{k,z} = \{ |\varepsilon_{k,z}| > \sqrt{2(1+\delta) \log \lambda_{k,z}} \} \quad \text{et} \quad R_{n,z} = \bigcup_{k=1}^n A_{k,z}$$

où $\lambda_{k,z}$ sera fixé plus tard. On a alors pour tout $\varepsilon > 0$

$$(3.9) \quad P(\tilde{h}_{n,z} > \varepsilon g(z)z^{-\frac{1}{2}}) \leq P(\{ \tilde{h}_{n,z} > \varepsilon g(z)z^{-\frac{1}{2}} \} \cap \bar{R}_{n,z}) + P(R_{n,z})$$

Appliquant le résultat de Heyde [10] sur la borne de Berry-Esséen on a

$$(3.10) \quad P(R_{n,z}) \leq \sum_{k=1}^n 2\phi(-\sqrt{2(1+\delta) \log \lambda_{k,z}}) + C\mu_3\sigma^{-3}E(Z_{k-1,z})^{-\frac{1}{2}}$$

$$\text{avec } \phi(x) = (2\pi)^{-\frac{1}{2}} \int_{-\infty}^x e^{-\frac{t^2}{2}} dt \leq (|x| \sqrt{2\pi})^{-1} e^{-\frac{x^2}{2}} \quad x < 0.$$

En remarquant que :

$$\gamma = \sum_{k=1}^{\infty} k^{-1} p_k \leq 1 - \sum_{k=2}^{+\infty} \left(1 - \frac{1}{k}\right) a_k = \gamma_0$$

la formule 4 du théorème 2 de [11] devient

$$(3.11) \quad \mathbb{E}(Z_{k-1,z})^{-\frac{1}{2}} \leq \gamma_0^{\frac{k-1}{2}} z^{-\frac{1}{2}}$$

Choissant $\lambda_{k,z} = k + f(z)$ avec $f(z) \rightarrow \infty$ on écrit

$$(3.12) \quad \mathbb{P}(\mathbf{R}_{n,z}) \leq [4\pi\delta^2(1+\delta)]^{-\frac{1}{2}} f(z)^{-\delta} [\log f(z)]^{-\frac{1}{2}} + C\mu_3\sigma^{-3}(1-\sqrt{\gamma_0})^{-1} z^{-\frac{1}{2}}$$

et donc quand $n, z \rightarrow \infty$, $\mathbb{P}(\mathbf{R}_{n,z}) \rightarrow 0$ uniformément dans \mathcal{P} .

De plus

$$(3.13) \quad \mathbb{P}(\{\tilde{h}_{n,z} > \varepsilon g(z)z^{-\frac{1}{2}}\} \cap \bar{\mathbf{R}}_{n,z}) \leq \sigma z^{\frac{1}{2}} \varepsilon^{-1} g(z)^{-1} \sum_{k=1}^n [k + f(z)]^{1+\delta} \mathbb{E}(Z_{k-1,z})^{-\frac{1}{2}} \\ \leq 2^\delta \sigma \varepsilon^{-1} g(z)^{-1} \left(\sum_{k=1}^{\infty} k^{1+\delta} \gamma_0^{\frac{k-1}{2}} \right) (1 + f(z))^{1+\delta}$$

si on choisit δ et f tels que $f(z)^{1+\delta} g(z)^{-1} \rightarrow 0$, on obtient le résultat.

COROLLAIRE 1. — Soient $c_1, c_2 > 0$ et $m'_{n,z}, \sigma'^2_{n,z}$ vérifiant

$$(3.14) \quad |m'_{n,z} - m| \leq c_1 \sqrt{\frac{m-1}{(m^n-1)z}}, \quad |\sigma'^2_{n,z} - \sigma^2| \leq c_2 \sqrt{\frac{2\sigma^4}{n}}$$

alors $h'_{n,z}$ obtenu en remplaçant m et σ^2 par $m'_{n,z}$ et $\sigma'^2_{n,z}$ dans $h_{n,z}$ est encore un $o_{\mathbb{P}}(g(z)z^{-\frac{1}{2}})$.

Démonstration. — Soit $\mathbf{B}_{k,z}$ obtenu en remplaçant m et σ^2 dans $\mathbf{A}_{k,z}$ et

$$(3.15) \quad \mathbf{C}_{k,z} = \left\{ |\varepsilon_{k,z}| > \left(1 + \frac{\delta}{2}\right) \sigma'_{n,z} \sqrt{2 \log \lambda_{k,z}} \right\} \\ \mathbf{D}_{k,z} = \left\{ |m'_{n,z} - m| (Z_{k-1,z})^{\frac{1}{2}} > \frac{\delta}{2} \sigma'_{n,z} \sqrt{2 \log \lambda_{k,z}} \right\}$$

On raisonne sur $\mathbf{C}_{k,z}$ comme précédemment. De plus

$$\mathbb{P}(\mathbf{D}_{k,z}) \leq \frac{c_1^2(m-1)m^{k-1}}{2\sigma'^2_{n,z}(m^n-1) \log f(z)\delta^2} \quad \text{et} \quad \sum_{k=1}^n \mathbb{P}(\mathbf{D}_{k,z}) \leq \frac{c_1^2(m-1)}{2\delta^2\sigma'^2_{n,z} \log f(z)}$$

Ce corollaire sera utile dans la 5^e partie.

Démonstration de la proposition 2. — Soit φ la fonction caractéristique de p et

$$\psi_a(t) = \exp\left(-\frac{it ma}{\sigma\sqrt{a}}\right) \left[\varphi\left(\frac{t}{\sigma\sqrt{a}}\right)\right]^a.$$

On a $|\Delta_{k,z}| \leq J_1 + J_2 + J_3$ où

$$J_1 = \int_{|t| \leq \frac{\sigma^3\sqrt{a}}{4\mu_3}} |\psi_a(t) - e^{-\frac{t^2}{2}}| dt$$

$$J_2 = \int_{\frac{\sigma^3\sqrt{a}}{4\mu_3} \leq |t| \leq \pi\sigma\sqrt{a}} |\psi_a(t)| dt \quad \text{et} \quad J_3 = \int_{|t| > \frac{\sigma^3\sqrt{a}}{4\mu_3}} e^{-\frac{t^2}{2}} dt$$

Or Petrov [15] a montré que :

$$(3.16) \quad J_1 \leq \frac{16\mu_3}{\sigma^3\sqrt{a}} \int |t|^3 e^{-\frac{t^2}{3}} dt \quad \text{et} \quad J_3 \leq \frac{8\mu_3}{\sigma^3\sqrt{a}}$$

qui donnent des majorations uniformes dans \mathcal{P} . De plus

$$J_2 = \sigma\sqrt{a} \int_{\frac{\sigma^2}{4\mu_3} < |t| < \pi} |\varphi(t)|^a dt.$$

Mais

$$|\varphi(t)|^2 = 1 - 2 \sum_{k \neq j} p_j p_k \left(\sin \frac{(k-j)t}{2}\right)^2 \leq 1 - 2 \sum_{k \neq j} a_j a_k \left(\sin \frac{(k-j)t}{2}\right)^2$$

$$\leq 1 - s^2(1 - |\varphi_0(t)|^2)$$

où on a posé $s = \sum a_k$ et $\varphi_0(t) = \sum \frac{a_k}{s} e^{ikt}$.

Or (2.2) entraîne que pour $t \neq 0 \pmod{2\pi}$ $|\varphi_0(t)| < 1$. Donc si on pose

$$0 < r_0 = \inf_{p \in \mathcal{P}} \frac{\sigma^2}{4\mu_3} \quad \text{et} \quad 0 < K = \min_{r_0 \leq t \leq \pi} -2 \log [1 - s^2(1 - |\varphi_0(t)|^2)]$$

on a

$$J_2 \leq 2\pi\sigma\sqrt{a} e^{-Kr_0}$$

ce qui conduit au résultat annoncé.

4. ESTIMATEURS ISSUS DE \hat{L} : CONSISTANCE ET NORMALITÉ CONJOINTE

On a d'après la formule (3.2)

$$(4.1) \quad \begin{aligned} \frac{\partial \hat{L}_{n,z}}{\partial m} &= \sigma^{-2} \sum_{k=1}^n (Z_{k,z} - mZ_{k-1,z}) \\ \frac{\partial \hat{L}_{n,z}}{\partial \sigma^2} &= -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4} \sum_{k=1}^n (Z_{k,z} - mZ_{k-1,z})^2 (Z_{k-1,z})^{-1} \end{aligned}$$

En annulant on retrouve les estimateurs cités dans l'introduction :

$$(4.2) \quad \begin{aligned} \hat{m}_{n,z} &= \left(\sum_{k=1}^n Z_{k,z} \right) \left(\sum_{k=1}^n Z_{k-1,z} \right)^{-1} \\ \hat{\sigma}_{n,z}^2 &= n^{-1} \sum_{k=1}^n (Z_{k,z} - \hat{m}_{n,z} Z_{k-1,z})^2 (Z_{k-1,z})^{-1} \end{aligned}$$

On notera aussi

$$\tilde{\sigma}_{n,z}^2 = n^{-1} \sum_{k=1}^n (Z_{k,z} - mZ_{k-1,z})^2 (Z_{k-1,z})^{-1}$$

qui estime σ^2 sans biais à m connu.

4.1. Consistance.

On sait que $\hat{m}_{n,z} \xrightarrow{P} m$ quand $n, z \rightarrow \infty$ ([18]). On a besoin de montrer le même type de convergence pour $\hat{\sigma}_{n,z}^2$.

PROPOSITION 3.

$$(4.3) \quad \hat{\sigma}_{n,z}^2 \xrightarrow{P} \sigma^2 \quad \text{quand} \quad n, z \rightarrow \infty$$

Démonstration. — On sait que le résultat est vrai à z fixé ([4])

$$(4.4) \quad \tilde{\sigma}_{n,z}^2 - \hat{\sigma}_{n,z}^2 = n^{-1} \left(\sum_{k=1}^n Z_{k-1,z} \right) (\hat{m}_{n,z} - m)^2$$

$$\text{et} \quad (\hat{m}_{n,z} - m)^2 = \left(\sum_{k=1}^n Z_{k-1,z} \right)^{-2} \left[\sum_{i,j} (Y(i,j) - m) \right]^2$$

où la sommation s'effectue pour $1 \leq j \leq 1 + Z_1^i + \dots + Z_{n-1}^i$ et $1 \leq i \leq z$.

Or ([18])

$$(4.5) \quad \frac{m-1}{(m^n-1)z} \sum_{k=1}^n Z_{k-1,z} \xrightarrow{P} 1 \quad n, z \rightarrow \infty$$

et $E \left[\sum_{i,j} (Y(i,j) - m) \right]^2 = \sigma^2 \frac{m^n - 1}{(m-1)z}$ entraînent

$$(4.6) \quad \tilde{\sigma}_{n,z}^2 - \hat{\sigma}_{n,z}^2 = O_p(n^{-1}) \quad n, z \rightarrow \infty$$

Mais ([4])

$$\text{Var } \tilde{\sigma}_{n,z}^2 = n^{-1} \left[2\sigma^4 + n^{-1} [E(Y-m)^4 - 3\sigma^4] \left(\sum_{k=0}^n E(Z_{k,z})^{-1} \right) \right]$$

c'est-à-dire que quand n et $z \rightarrow \infty$

$$\text{Var } \tilde{\sigma}_{n,z}^2 = o(n^{-1}), \quad \tilde{\sigma}_{n,z}^2 \xrightarrow{P} \sigma^2 \quad \text{et donc} \quad \hat{\sigma}_{n,z}^2 \xrightarrow{P} \sigma^2.$$

4.2. Normalité asymptotique de $\hat{\theta}_{n,z} = (\hat{m}_{n,z}, \hat{\sigma}_{n,z}^2)$.

(4.1) s'écrit :

$$(4.7) \quad \frac{\partial \hat{L}_{n,z}}{\partial m} = \sigma^{-2} \left(\sum_{k=1}^n Z_{k-1,z} \right) (\hat{m}_{n,z} - m) \quad \frac{\partial \hat{L}_{n,z}}{\partial \sigma^2} = \frac{n}{2\sigma^4} (\tilde{\sigma}_{n,z}^2 - \sigma^2)$$

A z fixé, le processus $\left(\frac{\partial \hat{L}_{n,z}}{\partial m}, \frac{\partial \hat{L}_{n,z}}{\partial \sigma^2} \right)$ est une martingale par rapport aux $\mathcal{F}_{n,z}$. Les théorèmes centraux limites de Scott [16] et Brown [1] rendent bien compte de la normalité asymptotique des marginales à z fixé mais les normalisations très différentes ($\sqrt{m^n z}$ et \sqrt{n}) rendent difficile une étude conjointe par le même type de méthodes.

On va plutôt utiliser \hat{L} .

PROPOSITION 4. — Si on pose $M_{n,z} = \sigma^{-1} \left[\frac{(m^n - 1)}{m - 1} z \right]^{\frac{1}{2}} (\hat{m}_{n,z} - m)$ et $S_{n,z} = \left(\frac{n}{2\sigma^4} \right)^{\frac{1}{2}} (\tilde{\sigma}_{n,z}^2 - \sigma^2)$ alors

$$(4.8) \quad (M_{n,z}, S_{n,z}) \xrightarrow{\mathcal{L}} \mathcal{N}(0, I_2) \quad \text{quand} \quad n, z \rightarrow \infty$$

Démonstration. — On sait déjà (Yanev [18]) que $M_{n,z} \xrightarrow{\mathcal{L}} \mathcal{N}(0, 1)$, $n, z \rightarrow \infty$. On va montrer que pour $t \in \mathbb{R}$ et $u \in \mathbb{R}^-$ fixés :

$$(4.9) \quad \Phi_{n,z}(t, u) = E \exp(it M_{n,z} + u S_{n,z}) \rightarrow \exp \frac{u^2 - t^2}{2}$$

Pour cela, on construit un nouveau processus de ramification dont la loi de reproduction a même espérance m que celui qu'on étudie et dont la variance est $\sigma^{*2} = \sigma^2 \left(1 - \frac{2u}{\sqrt{2n}}\right)^{-1} < \sigma^2$. Par exemple, soit $k \geq 1$ tel que $p_k p_{k+2} \neq 0$ (possible car $p \in \mathcal{P}$), on définit

$$(4.10) \quad \begin{aligned} q_i &= p_i & \text{si } i \neq k, k+1, k+2 \\ q_k &= p_k - \frac{1}{2}(\sigma^2 - \sigma^{*2}) & q_{k+1} = p_{k+1} + (\sigma^2 - \sigma^{*2}) \\ q_{k+2} &= p_{k+2} - \frac{1}{2}(\sigma^2 - \sigma^{*2}). \end{aligned}$$

Pour n assez grand, $q \in \mathcal{P}$. Soit $P^* = \bigotimes_{i,j=1}^7 q$ (cf. § 2) et $h_{n,z}^*$ le reste de la vraisemblance approchée du processus correspondant, et enfin $H_{n,z} = h_{n,z} - h_{n,z}^*$

$$(4.11) \quad \begin{aligned} \Phi_{n,z}(t, u) &= E^* \left[\exp(it M_{n,z} + u S_{n,z}) \frac{P}{P^*}(Z_{0,z}, Z_{1,z}, \dots, Z_{n,z}) \right] \\ P(Z_{0,z}, \dots, Z_{n-1,z}) &= (\sigma \sqrt{2\pi})^{-n} (Z_{0,z} \dots Z_{n-1,z})^{-\frac{1}{2}} \exp\left(-\frac{1}{2} n \frac{\tilde{\sigma}_{n,z}^2}{\sigma^2} + h_{n,z}\right) \\ P^*(Z_{0,z}, \dots, Z_{n-1,z}) &= (\sigma^* \sqrt{2\pi})^{-n} (Z_{0,z} \dots Z_{n-1,z})^{-\frac{1}{2}} \exp\left(-\frac{1}{2} n \frac{\tilde{\sigma}_{n,z}^2}{\sigma^{*2}} + h_{n,z}^*\right) \end{aligned}$$

d'où

$$\Phi_{n,z}(t, u) = V_n E^* \exp(it M_{n,z} + H_{n,z})$$

où on a posé $V_n = \left(1 - \frac{2u}{\sqrt{2n}}\right)^{-\frac{n}{2}} \exp -u \sqrt{\frac{n}{2}}$ qui tend vers $e^{\frac{u^2}{2}}$ quand $n \rightarrow \infty$.

Si $E_1 = E^*[\exp(it M_{n,z} + H_{n,z})]$ on a

$$(4.12) \quad |E_1 - e^{-\frac{t^2}{2}}| \leq |E^*(\exp it M_{n,z}) - e^{-\frac{t^2}{2}}| + |E^*(\exp H_{n,z}) - 1|$$

Mais pour $\alpha > 0$, on peut écrire

$$(4.13) \quad E^* |1 - \exp H_{n,z}| \leq e^\alpha - 1 + E^*[(\exp H_{n,z}) 1_{|H_{n,z}| > \alpha}] + P^*(|H_{n,z}| > \alpha)$$

et

$$(4.14) \quad \begin{aligned} E^*[(\exp H_{n,z}) 1_{|H_{n,z}| > \alpha}] &\leq \left(E^* \left[(\exp 2H_{n,z}) \frac{P^*}{P}(Z_{0,z}, \dots, Z_{n,z}) \right] \right)^{\frac{1}{2}} \\ &\times \left(E^* \left[1_{|H_{n,z}| > \alpha} \frac{P}{P^*}(Z_{0,z}, \dots, Z_{n,z}) \right] \right)^{\frac{1}{2}} \leq V_n^{-1} (E(\exp 2u S_{n,z}))^{\frac{1}{2}} P(|H_{n,z}| > \alpha)^{\frac{1}{2}} \end{aligned}$$

D'après [8], $S_{n,1} \xrightarrow{\mathcal{L}} \mathcal{N}(0, 1)$ quand $n \rightarrow \infty$. La démonstration s'adapte très facilement au cas $S_{n,z}$, $n, z \rightarrow \infty$, donc $(E(\exp 2uS_{n,z}))^{\frac{1}{2}} \rightarrow e^{u^2}$ quand $n, z \rightarrow \infty$. De plus pour tout α , $P(|H_{n,z}| > \alpha)$ et $P^*(|H_{n,z}| > \alpha)$ tendent vers 0 (cf. § 3), d'où

$$(4.15) \quad \overline{\lim}_{n,z} |E_1 - e^{-\frac{t^2}{2}}| \leq \overline{\lim}_{n,z} |E^*(\exp it M_{n,z}) - e^{-\frac{t^2}{2}}|$$

ce qui, en utilisant [18] qui reste vrai sous P^* permet de conclure.

COROLLAIRE 2. — Si on remplace m par $\hat{m}_{n,z}$ dans $S_{n,z}$, le résultat précédent reste vrai.

Démonstration. — Ceci résulte de $(\tilde{\sigma}_{n,z}^2 - \hat{\sigma}_{n,z}^2) = O_p(n^{-1})$, $n, z \rightarrow \infty$ (prop. 3).

5. EFFICACITÉ ASYMPTOTIQUE DE $\hat{\theta}_{n,z}$

Dans les processus stochastiques dépendant d'un seul paramètre (fonction de m pour un Galton-Watson), à vraisemblance assez régulière, Heyde [9] a montré l'existence d'un estimateur du maximum de vraisemblance et étudié ses propriétés d'efficacité. On utilisera ici les mêmes outils mais pour la vraisemblance approchée \hat{L} et pour le paramètre $\theta = (m, \sigma^2)$. Posons :

$$(5.1) \quad u_{k,z}^1 = \frac{\partial \hat{L}_{k,z}}{\partial m} - \frac{\partial \hat{L}_{k-1,z}}{\partial m}, \quad u_{k,z}^2 = \frac{\partial \hat{L}_{k,z}}{\partial \sigma^2} - \frac{\partial \hat{L}_{k-1,z}}{\partial \sigma^2}$$

$$K_{k,z}(\theta) = \begin{pmatrix} \frac{\partial u_{k,z}^1}{\partial m} & \frac{\partial u_{k,z}^1}{\partial \sigma^2} \\ \frac{\partial u_{k,z}^2}{\partial m} & \frac{\partial u_{k,z}^2}{\partial \sigma^2} \end{pmatrix}$$

$$(5.2) \quad J_{n,z}(\theta) = \sum_{k=1}^n K_{k,z}(\theta)$$

$$I_{n,z}(\theta) = - \sum_{k=1}^n E[K_{k,z}(\theta) | \mathcal{F}_{k-1,z}] = \begin{pmatrix} \sigma^{-2} \sum_{k=1}^n Z_{k-1,z} & 0 \\ 0 & \frac{n}{2\sigma^4} \end{pmatrix}$$

$I_{n,z}(\theta)$ joue le rôle d'information de Fisher, mais il est à remarquer que

$$I_{n,z}(\theta) \neq \left(\sum_{k=1}^n E(u_{k,z}^i u_{k,z}^j | \mathcal{F}_{k-1,z}) \right)_{\substack{i=1,2 \\ j=1,2}}$$

car \hat{L} n'est pas une véritable vraisemblance.

On a

$$EI_{n,z}(\theta) = \begin{pmatrix} \frac{z(m^n - 1)}{\sigma^2(m - 1)} & 0 \\ 0 & \frac{n}{2\sigma^4} \end{pmatrix}.$$

LEMME 1.

$$(5.3) \quad I_{n,z}(\theta)[EI_{n,z}(\theta)]^{-1} \xrightarrow{P} I_2, \quad n, z \rightarrow \infty$$

c'est une conséquence triviale de [18]. Au contraire à z fixé la limite est aléatoire [9]; $z \rightarrow \infty$ fait jouer la loi des grands nombres.

PROPOSITION 5. — Soit $C = (c_1, c_2)$ avec c_1, c_2 fixés > 0 . Soit $p^{n,z} \in \mathcal{P}$ tel que le θ associé soit $\theta_{n,z} = \theta + 2C[EI_{n,z}(\theta)]^{-\frac{1}{2}}$, $P^{n,z} = \bigotimes_{i,j=1}^2 p^{n,z}$ et soit $\Gamma = [-c_1, c_1] \times [-c_2, c_2]$, alors

$$i) \lim_{n,z} \{ P[(\hat{\theta}_{n,z} - \theta)[EI_{n,z}(\theta)]^{\frac{1}{2}} \in \Gamma] - P^{n,z}[(\hat{\theta}_{n,z} - \theta_{n,z})[EI_{n,z}(\theta)]^{\frac{1}{2}} \in \Gamma] \} = 0.$$

ii) Si $T_{n,z}$ est un autre estimateur de θ vérifiant i) alors

$$\overline{\lim}_{n,z} P[(T_{n,z} - \theta)[EI_{n,z}(\theta)]^{\frac{1}{2}} \in \Gamma] \leq \lim_{n,z} P[(\hat{\theta}_{n,z} - \theta)[EI_{n,z}(\theta)]^{\frac{1}{2}} \in \Gamma].$$

iii) Si $T_{n,z}$ en plus de i) est tel que

$$(T_{n,z} - \theta)[EI_{n,z}(\theta)]^{\frac{1}{2}} \xrightarrow{\mathcal{L}} \mathcal{N}(0, \gamma)$$

alors $\gamma \geq I_2$ au sens de l'ordre sur les matrices symétriques.

Démonstration. — i) Le corollaire 2 entraîne

$$P[(\hat{\theta}_{n,z} - \theta)[EI_{n,z}(\theta)]^{\frac{1}{2}} \in \Gamma] \rightarrow \int_{\Gamma} \frac{e^{-\frac{1}{2}(x^2+y^2)}}{2\pi} dx dy.$$

De plus vu la définition de \mathcal{P} , ceci est vrai uniformément et

$$(5.4) \quad [EI_{n,z}(\theta)]^{-1} E^{n,z} I_{n,z}(\theta_{n,z}) \rightarrow I_2 \quad n, z \rightarrow \infty.$$

ii) On reprend le raisonnement de Weiss-Wolfowitz [13] pour un paramètre de \mathbb{R}^2 mais avec les vraisemblances approchées.

Si on pose

$$(5.5) \quad \Lambda_{n,z} = \log \frac{\mathbf{P}^{n,z}}{\mathbf{P}}(Z_{0,z}, \dots, Z_{n,z})$$

on a

$$\Lambda_{n,z} = \hat{\Lambda}_{n,z} + H_{n,z}$$

avec $\hat{\Lambda}_{n,z} = \hat{L}_{n,z}(\theta_{n,z}) - \hat{L}_{n,z}(\theta)$ et $H_{n,z} = o_p\left(\frac{g(z)}{\sqrt{z}}\right)$ (prop. 2 et cor. 1). Mais

$$(5.6) \quad \hat{\Lambda}_{n,z} = \left(\sum_{k=1}^n u_{k,z}(\theta) \right) (\theta_{n,z} - \theta)' + \frac{1}{2} (\theta_{n,z} - \theta) J_n(\theta_{n,z}^*) (\theta_{n,z} - \theta)'$$

avec

$$\theta_{n,z}^* = (m + \lambda_1^*(m_{n,z} - m), \sigma^2 + \lambda_2^*(\sigma_{n,z}^2 - \sigma^2)) \quad \lambda_1^*, \lambda_2^* \in [0, 1].$$

De plus pour tout $\bar{\theta}$

$$(5.7) \quad \frac{d}{d\bar{\theta}} \hat{L}_{n,z}(\bar{\theta}) = \sum_1^n u_{k,z}(\bar{\theta}) = \sum_{k=1}^n u_{k,z}(\theta) + (\bar{\theta} - \theta) J_{n,z}(\theta^{**})$$

avec $\theta^{**} = (m + \lambda_1^{**}(\bar{m} - m), \sigma^2 + \lambda_2^{**}(\bar{\sigma}^2 - \sigma^2))$, $\lambda_1^{**}, \lambda_2^{**} \in [0, 1]$.

Faisant $\bar{\theta} = \hat{\theta}_{n,z}$, (5.7) entraîne

$$(5.8) \quad \begin{aligned} \hat{\Lambda}_{n,z} &= -(\hat{\theta}_{n,z} - \theta) J_{n,z}(\theta_{n,z}^{**}) (\theta_{n,z} - \theta)' + \frac{1}{2} (\theta_{n,z} - \theta) J_{n,z}(\theta_{n,z}^*) (\theta_{n,z} - \theta)' \\ &= -2(\hat{\theta}_{n,z} - \theta) J_{n,z}(\theta_{n,z}^{**}) [E I_{n,z}(\theta)]^{-\frac{1}{2}} C' + 2C [E I_{n,z}(\theta)]^{-\frac{1}{2}} J_{n,z}(\theta_{n,z}^*) [E I_{n,z}(\theta)]^{-\frac{1}{2}} C' \end{aligned}$$

A ce stade nous avons besoin du lemme de calcul suivant (admis)

LEMME 2. — 1)

$$(5.9) \quad [E I_{n,z}(\theta)]^{-\frac{1}{2}} J_{n,z}(\theta) [E I_{n,z}(\theta)]^{-\frac{1}{2}} \xrightarrow{\mathbf{P}} -I_2 \quad n, z \rightarrow \infty$$

2) Si $\bar{\theta}_{n,z} = \theta + D_{n,z} [E I_{n,z}(\theta)]^{-\frac{1}{2}}$ avec $D_{n,z}$ vecteur borné ($n, z \rightarrow \infty$) alors

$$(5.10) \quad E [I_{n,z}(\theta)]^{-\frac{1}{2}} [J_{n,z}(\bar{\theta}_{n,z}) - J_{n,z}(\theta)] [E I_{n,z}(\theta)]^{-\frac{1}{2}} = O_p(n^{-\frac{1}{2}}).$$

Ceci entraîne :

$$(5.11) \quad \hat{\Lambda}_{n,z} = -2CC' + 2(\hat{\theta}_{n,z} - \theta) [E I_{n,z}(\theta)]^{\frac{1}{2}} C' + O_p(n^{-\frac{1}{2}})$$

et donc

$$(5.12) \quad \Lambda_{n,z} = -2CC' + 2(\hat{\theta}_{n,z} - \theta)[E I_{n,z}(\theta)]^{\frac{1}{2}}C' + O_P(n^{-\frac{1}{2}}) + o_P\left(\frac{g(z)}{\sqrt{z}}\right)$$

La suite est classique (Weiss-Wolfowitz [17]).

iii) est une conséquence de ii) lorsque $T_{n,z}$ est un compétiteur asymptotiquement normal.

Remarques. — 1) Ayant terminé ce travail, nous avons eu connaissance de Eschenbach-Winkler ([7]). Ils proposent la même approximation de la vraisemblance mais n'en n'évaluent pas la précision.

2) N. Keiding nous a signalé une analogie avec l'estimation dans les processus de ramification-diffusion ([2]) :

$$\text{soit} \quad \theta_z = (1 + \beta z^{-1}, \sigma^2), \quad \beta > 0 \quad \text{et} \quad X_{t,z} = z^{-1}Z_{[tz],z}$$

alors quand $z \rightarrow \infty$, $X_{t,z} \xrightarrow{\mathcal{L}} X_t$ où X_t vérifie

$$dX_t = \beta X_t dt + \sigma \sqrt{X_t} dW_t, \quad X_0 = 1$$

(W_t brownien standard).

Cette équation est l'analogie de (1.1) mais le problème théorique de l'estimation de (β, σ) au vue de X est un peu différent. La méthode naturelle consiste à estimer d'abord σ^2 par variation quadratique puis β par

$$(X_t - 1) \left[\int_0^t X_s ds \right]^{-1}.$$

BIBLIOGRAPHIE

- [1] B. M. BROWN, Martingale central limit theorems, *Ann. Math. Stat.*, t. **42**, 1971, p. 59-66.
- [2] B. M. BROWN, J. I. HEWITT, Inference for the diffusion branching process, *J. A. P.*, t. **12**, 1975, p. 588-594.
- [3] J. P. DION, *Estimation des probabilités initiales et de la moyenne d'un processus de Galton-Watson*. Thèse, Montréal, 1972.
- [4] J. P. DION, Estimation of the variance of a branching process. *Ann. Stat.*, t. **3**, 1975, p. 1183-1187.
- [5] J. DION, N. KEIDING, Statistical inference in branching processes, In *Branching Processes* edited by Joffe and Ney. *Advances in Probability and Related Topics*, t. **5**, 1978, M. Dekker.
- [6] C. DUBY, A. ROUAULT, *C. R. A. S.*, t. **290**, 1980, p. 339-341.
- [7] W. ESCHENBACH, W. WINKLER, Maximum-likelihood-shätzungen beim Verzweigungsprozess von Galton-Watson. *Math. Operationsforsch. u. Stat.*, t. **6**, 1975, p. 213-224.
- [8] C. C. HEYDE, On estimating the variance of the offspring distribution in a simple branching process, *Adv. in Appl. Prob.*, t. **6**, 1974, p. 421-433.

- [9] C. C. HEYDE, On an optimal asymptotic property of the maximum likelihood estimator of a parameter from a stochastic process. *Stochastic Processes and their applications*, t. **8**, 1978, p. 1-9.
- [10] C. C. HEYDE, Some almost sure convergence theorems for branching processes. *Z. Wahrscheinlichkeitstheorie*, t. **20**, 1971, p. 189-192.
- [11] C. C. HEYDE, B. M. BROWN, An invariance principle and some convergence rate results for branching processes. *Z. Wahrscheinlichkeitstheorie*, t. **20**, 1971, p. 271-278.
- [12] C. C. HEYDE, P. D. FEIGIN, On efficiency and exponential families in stochastic process estimation. In : *Statistical distributions in scientific work* (ed. G. P. Patil, S. Kotz, J. K. Ord), Reidel, Dordrecht, t. **1**, 1975, p. 227-240.
- [13] P. JAGERS, *Branching processes with biological applications*, Wiley, 1975.
- [14] N. KEIDING, S. LAURITZEN, Marginal maximum likelihood estimates and estimation of the offspring mean in a branching process. *Scand. J. Statist.*, t. **5**, 1978.
- [15] PETROV, *Sums of independent random variables*, Springer-Verlag, 1975.
- [16] SCOTT, A central limit theorem for martingales and an application to branching processes. *Stochastic Processes and Their Applications*, t. **6**, 1978, p. 241-252.
- [17] L. WEISS, J. WOLFOWITZ, Generalized maximum likelihood estimators. *Theory of Prob. and its Appl.*, vol. **XI**, 1966, p. 58-80.
- [18] N. M. YANEV, On the statistics of branching processes. *Theory of Prob. and its appl.*, t. **20**, 1975, p. 612-622.

(Manuscrit reçu le 10 septembre 1981)

B. Quelques rappels et commentaires sur les dualités de Young et

Legendre

QUELQUES RAPPELS ET COMMENTAIRES SUR LES DUALITES DE
YOUNG ET LEGENDRE.

Afin de faciliter la lecture des chapitres qui suivent, où les notions de dualité sont d'un usage constant, nous en rappelons les définitions et principales propriétés, sans prétendre ni à l'originalité ni à la généralité (pour lesquelles on peut se reporter à [1], [3], [4]). Puis nous précisons le cadre probabiliste (transformée de Cramer) dans lequel elles interviendront.

Dualité de Young.

Soit V un espace de Banach de dual V^* et $\bar{\mathbb{R}} = \mathbb{R} \cup \{+\infty\} \cup \{-\infty\}$. On note $\Gamma(V)$ l'ensemble des fonctions f convexes s.c.i. de V dans $\bar{\mathbb{R}}$ telles que si $f = -\infty$ en un point $f = -\infty$ partout, et $\Gamma_0(V)$ le sous-ensemble de $\Gamma(V)$ obtenu en ôtant les constantes $+\infty$ et $-\infty$.

Pour toute f de V dans $\bar{\mathbb{R}}$, on définit la conjuguée (ou duale) f^* de f , de V^* dans $\bar{\mathbb{R}}$ par

$$f^*(v^*) = \sup_{v \in V} \langle v, v^* \rangle - f(v)$$

f^* est finie partout si et seulement si f vérifie la condition de croissance

$$(CC) \quad \frac{f(v)}{|v|} \xrightarrow{|v| \rightarrow +\infty} +\infty$$

$f^{**} = f$ si et seulement si f est dans $\Gamma(V)$.

La dualité établit une bijection entre $\Gamma_0(V)$ et $\Gamma_0(V^*)$.

Dualité de Legendre.

Soit f une fonction différentiable sur un ouvert C de \mathbb{R}^k . Si f' est une bijection de C sur $D \subset \mathbb{R}^k$, on appelle dual de Legendre du couple (C, f) le couple (D, g) où g , fonction de D dans \mathbb{R} est définie par :

$$\begin{aligned} g(x) &= \langle \theta, x \rangle - f(\theta) & x \in D \\ f'(\theta) &= x & \theta \in C \end{aligned}$$

En particulier, si f est de classe C^1 sur \mathbb{R} et f' strictement croissante de \mathbb{R} sur $D \subset \mathbb{R}$, g est de classe C^1 sur D ouvert et on a

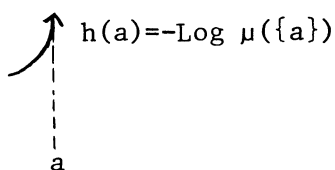
$$g' \circ f' = \text{Id}_{\mathbb{R}} \quad f' \circ g' = \text{Id}_D .$$

Transformée de Cramer.

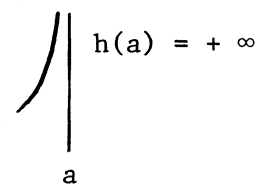
Soit μ une mesure de Radon positive sur \mathbb{R} . Le plus grand ouvert de finitude de la fonction convexe $\hat{\mu}$ de \mathbb{R} dans $\bar{\mathbb{R}}$ définie par (transformée de Laplace) :

$$\hat{\mu}(\theta) = \int_{\mathbb{R}} e^{\theta x} \mu(dx)$$

est un intervalle I . On suppose que I contient 0 ce qui entraîne que $L = \text{Log } \hat{\mu}$ est strictement convexe et analytique dans I . On note h la duale de Young de L , le couple $(h, L'(I))$ étant dual de Legendre de (L, I) . h s'appelle la transformée de Cramer de μ . Elle est s.c.i., finie à l'intérieur de l'enveloppe convexe S du support de μ et infinie à l'extérieur. Le comportement de L au bord de I affecte la stricte convexité de h ; le comportement de h aux 2 bords de S est en rapport avec la présence ou non de masses en ces points, représenté par les schémas suivants (bord droit)



$$\mu(\{a\}) > 0, \mu(]a, +\infty[) = 0$$



$$\mu(\{a\}) = 0, \mu(]-\infty, a[) > 0, \mu(]a, +\infty[) = 0$$

Dans le cas où une seule fonction L intervient (Chap. C, D, E) on peut détailler les propriétés du couple L, h aux bords ou même travailler dans \mathbb{R}^2 (pour Chap. D). Comme dans le modèle markovien (Chap. F, G, H) on manipule des familles de mesures et que les problèmes de bords risquaient de cacher la structure du phénomène, on a supposé que L et h étaient définies partout. ([2] rend bien compte du comportement d'une marche markovienne avec masses aux bords). Nos résultats sont aisément généralisables au cas d'un support compact fixe pour toutes les mesures en jeu, sans masse au bord. La famille de fonctions h constituant l'outil essentiel, c'est sur h - qui contient toute l'information sur μ - que seront écrites nos autres hypothèses, ce qui permettra à l'occasion (Chap. G, § I) d'autonomiser certains problèmes (non probabilistes) en ne considérant pas nécessairement h comme une transformée de Cramer. On peut remarquer à ce propos qu'une certaine tradition - inconsciemment ironique ? - a fait noter L la Log-Laplace (qui a le statut d'un Hamiltonien) et h la transformée de Cramer (qui a le statut d'un Lagrangien) ...

Références :

- [1] AZENCOTT (R.), *Cours de l'Ecole d'Eté de Saint-Flour, Lecture Notes in Math, n° 774 (1978).*
- [2] AZENCOTT (R.), RUGET (G.), *Mélanges d'équations différentielles et grands écarts à la loi des grands nombres. Zeits. für Wahrs., 1977, t. 38, p. 1-54*
- [3] EKELAND (I.), TURNBULL (T.), *Infinite dimensional optimization and convexity. The University of Chicago Press, 1983.*
- [4] ROCKAFELLAR (R.T.), *Convex Analysis, Princeton University Press, 1970.*

C. Lois de Zipf et sources markoviennes.

Lois de Zipf et sources markoviennes

par

Alain ROUAULT

Département de Mathématiques, Université Paris-Sud,
Centre d'Orsay, 91405 Orsay

RÉSUMÉ. — On étudie ici les lois de Zipf — d'origine linguistique — dans le cadre d'un n -échantillon d'une v. a. dénombrable. Puis on montre que la suite des valeurs prises par une chaîne de Markov finie, avec un état ne se succédant pas à lui-même (blanc) peut être considérée comme un échantillon d'un ensemble de mots, vérifiant les lois de Zipf. On montre aussi un résultat asymptotique sur une v. a. Z_n^{m*} , nombre de mots — parmi les n premiers — ayant été utilisés au moins m fois.

SUMMARY. — We study here Zipf's laws (of linguistic origine) in the context of a n -sample of a denumerable random value. Then we prove that the sequence of values taken by a finite Markov chain with a not-self-following state (blank) may be considered as a sample from a set of words, fulfilling Zipf's laws. We also prove an asymptotical result on Z_n^{m*} , number of words (among the n first) used more than m -times.

I. INTRODUCTION

Le nom de Zipf est connu des linguistes, des démographes, des biologistes et une littérature mathématique substantielle a été consacrée à ses lois. Nous emploierons la terminologie linguistique pour introduire le problème. Il s'agit d'étudier le rapport entre les mots potentiels d'un texte (le vocabulaire de l'auteur) et les mots rencontrés effectivement au cours de sa lecture, les répétitions...

La première loi concerne l'étude, lorsque la lecture avance, du rapport

du nombre de mots rencontrés s fois, au nombre total de mots différents rencontrés $s = 1, 2, \dots$

La seconde concerne la relation entre la fréquence d'un mot, et la place qu'il occupe dans le rangement des mots par ordre de fréquences décroissantes.

Enfin une autre loi (sans nom) dit que pour n grand, le nombre de mots différents dans un texte de longueur n , considéré comme aléatoire a approximativement une distribution de Gauss.

Selon un premier point de vue [7] [8] un auteur écrit un texte en effectuant des tirages successifs de mots, avec remplacement, dans un dictionnaire (fini ou non), muni d'une loi de probabilité.

Selon un second [3] [4] [5], on opère uniformément dans un ensemble fini N divisé en M classes et on fait croître N et M .

Nous adoptons le premier. McNeil [8] suppose le dictionnaire fini et estime sa taille. Mandelbrot [7] fait des approximations qui prennent tout leur sens quand on les rapproche des résultats de Karlin [6]. Dans le III nous remarquons ainsi que la première loi de Zipf est une conséquence immédiate de ces résultats (sous une condition simple — grosso modo $p_k \simeq k^{-\frac{1}{\gamma}}$, $0 < \gamma < 1$ pour la probabilité). La seconde nous dira seulement que la k -ième fréquence (par ordre décroissant) converge p. s. vers la k -ième probabilité. Ceci est une simple conséquence de la loi forte des grands nombres.

Dans le IV, on part d'un alphabet fini, comprenant un élément « blanc » ou « espace » et on suppose que le texte est composé, lettre après lettre, selon un procédé markovien. Ceci implique pour les mots formés, un tirage avec remplacement et nous montrons que la probabilité correspondante vérifie la condition ci-dessus, le γ étant le paramètre malthusien d'un certain processus de ramification.

Dans V, nous démontrons une convergence de Z_n^{m*} (nombre de mots, parmi les n premiers, rencontrés au moins m fois) dans le cas où m et n tendent vers l'infini, le quotient $\frac{m}{n}$ restant entre certaines limites.

II. NOTATIONS

Considérons un ensemble dénombrable assimilé à \mathbb{N}^* muni d'une probabilité $p = (p_n)_{n \geq 1}$. Pour tout x tel que $0 < x < 1$, définissons

$$\alpha(x) = \text{card} \{ k : p_k > x \} \quad (2.1)$$

CONDITION 1. — $\forall n, p_n > 0$ et $\exists 0 < \gamma < 1$ et $L [0, 1] \rightarrow \mathbb{R}^+$ à variation lente au voisinage de 0, tels que

$$\alpha(x) = x^{-\gamma}L(x) \tag{2.2}$$

Soit $(X_n)_{n \geq 1}$ une suite de v. a. à valeurs dans \mathbb{N}^* , indépendantes et de même loi p . Si $\delta_n, n = 1, 2, \dots$ est la n -ième fonction de Kronecker, notons :

$$\begin{aligned} X_n^k &= \sum_{m=1}^n \delta_k(X_m) & k = 1, 2, \dots \\ Z_n^r &= \sum_{k \geq 1} \delta_r(X_n^k) & r = 1, 2, \dots \text{ (nombre de valeurs prises } r \text{ fois)} \\ Z_n^{r*} &= \sum_{s \geq r} Z_n^s, & Z_n^* = Z_n^{1*} \text{ (nombre de valeurs prises)} \end{aligned} \tag{2.3}$$

On notera $X_n^{(1)} \geq X_n^{(2)} \geq \dots \geq X_n^{(Z_n^*)} > 0$ le réarrangement par ordre décroissant des Z_n^* termes non nuls de la suite $(X_n^k)_{k \geq 1}$ et $p_{(1)} \geq \dots \geq p_{(r)} \geq \dots$ celui correspondant aux p .

III. LOIS DE ZIPF

PROPOSITION 1 (1^{re} loi). — Sous la condition 1, quand n tend vers l'infini, on a

$$\forall s \geq 1, \quad \lim \text{ p. s. } \frac{Z_n^s}{Z_n^*} = \frac{\gamma}{\Gamma(1-\gamma)} \frac{\Gamma(s-\gamma)}{\Gamma(s+1)} \tag{3.1}$$

Démonstration. — Karlin a démontré [6] :

$$\lim_{n \rightarrow +\infty} \text{ p. s. } \frac{Z_n^*}{EZ_n^*} = 1, \quad \forall s \geq 1 \quad \lim_{n \rightarrow +\infty} \text{ p. s. } \frac{Z_n^s}{EZ_n^s} = 1 \quad \text{sans condition}$$

$$EZ_n^* \underset{n \rightarrow +\infty}{\sim} \Gamma(1-\gamma)n^\gamma L\left(\frac{1}{n}\right), \quad EZ_n^s \underset{n \rightarrow +\infty}{\sim} \gamma \frac{\Gamma(s-\gamma)}{\Gamma(s+1)} n^\gamma L\left(\frac{1}{n}\right)$$

sous la condition 1

d'où le résultat. Quand s tend vers l'infini $\frac{\gamma}{\Gamma(1-\gamma)} \frac{\Gamma(s-\gamma)}{\Gamma(s+1)} \sim \frac{\gamma}{\Gamma(1-\gamma)} \frac{1}{s^{1+\gamma}}$ qui est de la forme proposée par différents auteurs [3] [4] [5] comme caractérisant la loi de Zipf.

Remarquons que si $s = 1$, on a $\lim \text{p. s. } \frac{Z_n^1}{Z_n^*} = \gamma$, d'où une possibilité d'estimer γ .

PROPOSITION 2 (2^e loi). — Quand n tend vers l'infini, on a

$$\forall r \geq 1 \quad \lim \text{p. s. } \frac{X_n^{(r)}}{n} = p_{(r)} \quad (3.2)$$

Démonstration. — C'est une conséquence immédiate de la loi forte des grands nombres. Nous supposons pour simplifier que $\forall n \ p_n \geq p_{n+1}$. Soit r fixé et

$$a = \alpha(p_r) \quad b = \alpha(p_r^-) \quad a < r \leq b \quad (3.3)$$

soit R tel que

$$\sum_{k=R+1}^{+\infty} p_k < p_{b+1} \quad (3.4)$$

soit $j_1 < j_2 < \dots < j_n < \dots$ l'image de la fonction α et

$$0 < \varepsilon < \frac{1}{2} \min (p_{j_s} - p_{j_{s+1}}) \quad s = 1, 2, \dots, \alpha(p_R^-) \quad (3.5)$$

Alors pour $s = 1, 2, \dots, \alpha(p_R^-)$ les intervalles $[p_{j_s} - \varepsilon, p_{j_s} + \varepsilon]$ ont deux à deux une intersection vide. Or

$$\text{p. s. } \exists N_1 \text{ tel que } n > N_1 \Rightarrow \forall k \leq R \quad \left| \frac{X_n^k}{n} - p_k \right| < \varepsilon$$

$$\text{p. s. } \exists N_2 \text{ tel que } n > N_2 \Rightarrow \left| \frac{\sum_{k=R+1}^{+\infty} X_n^k}{n} - \sum_{k=R+1}^{+\infty} p_k \right| < \varepsilon$$

donc

$$\text{p. s. } \exists N \text{ tel que } n > N \Rightarrow \forall k \leq b \quad \left| \frac{X_n^k}{n} - p_k \right| < \varepsilon$$

$$\forall b < k \leq R \quad \frac{X_n^k}{n} < p_k + \varepsilon \leq p_{b+1} + \varepsilon$$

$$\forall k \geq R + 1 \quad \frac{X_n^k}{n} \leq \frac{1}{n} \sum_{R+1}^{+\infty} X_n^j$$

$$< \left(\sum_{R+1}^{+\infty} p_j \right) + \varepsilon < p_{b+1} + \varepsilon$$

Par suite, p. s. il existe N , tel que $n > N$ entraîne :
il y a exactement

$$\begin{array}{rcl}
 j_1 & \text{fréquences dans l'intervalle } [p_{j_1} - \varepsilon, p_{j_1} + \varepsilon] & \\
 j_2 - j_1 & \text{»} & \text{»} \quad [p_{j_2} - \varepsilon, p_{j_2} + \varepsilon] \\
 \dots & \dots & \dots \\
 b - a & \text{»} & \text{»} \quad [p_b - \varepsilon, p_b + \varepsilon] = [p_r - \varepsilon, p_r + \varepsilon]
 \end{array}$$

et donc

$$\text{p. s. } \forall \varepsilon > 0 \quad \exists N : \quad n > N \Rightarrow \left| \frac{X_n^{(r)}}{n} - p_r \right| < \varepsilon.$$

D'où la proposition.

On remarque que de la majoration exponentielle classique (à r fixé) de $P\left(\left|\frac{X_n^{(r)}}{n} - p_r\right| > \varepsilon\right)$, on peut déduire une majoration exponentielle pour $P\left(\left|\frac{X_n^{(r)}}{n} - p_{(r)}\right| > \varepsilon\right)$.

IV. SOURCES MARKOVIENNES

1) Définitions

a) On appelle source markovienne une chaîne de Markov définie sur un espace $\bar{E} = \{e_1, \dots, e_d, \sigma\}$ ($d \geq 1$) ayant pour probabilité initiale la mesure de Dirac en σ et dont tous les éléments de la matrice de transition sont strictement positifs, sauf celui correspondant à un passage de σ à σ .

$\Omega = \bar{E}^{\mathbb{N}}$, $X = (X_k)_{k \in \mathbb{N}}$, les coordonnées, $P(X_0 = \sigma) = 1$

$$A = \begin{bmatrix} p_{11} & \dots & p_{1d} & p_{1\sigma} \\ p_{d1} & \dots & p_{dd} & p_{d\sigma} \\ p_{\sigma 1} & \dots & p_{\sigma d} & 0 \end{bmatrix} \quad (4.1)$$

$$T_1 = \inf \{n \geq 1, X_n = \sigma\} \quad T_{k+1} = \inf \{n > T_k, X_n = \sigma\} \quad k \geq 1$$

$$E = \{e_1, \dots, e_d\} \quad E^* = \bigcup_{n=1}^{+\infty} E^n$$

b) Un élément de E^n s'appelle un mot de longueur n et sera noté $x = x_1 x_2 \dots x_n$ avec, pour tout i , $x_i \in E$.

$$\forall k \quad \Omega \xrightarrow{M_k} E^* \quad \begin{array}{l} M_1(\omega) = X_1(\omega) X_2(\omega) \dots X_{T_1(\omega)-1}(\omega) \\ k > 1 \quad M_k(\omega) = X_{T_{k-1}(\omega)+1}(\omega) \dots X_{T_k(\omega)-1}(\omega) \end{array} \quad (4.2)$$

2) Échantillonnage

PROPOSITION 3. — Les M_k sont indépendants et de même loi (notée μ).

Ceci résulte du caractère markovien et irréductible de la chaîne. Les n premiers mots (aléatoires) sont donc n tirages indépendants dans l'ensemble E^* suivant la loi μ . On est alors conduit à se demander si la loi μ vérifie la condition 1 (cf. II) c'est-à-dire à étudier :

$$0 < t < 1 \quad \alpha(t) = \text{card} \{ x \in E^* ; \mu(x) > t \}$$

La réponse est donnée par le

THÉORÈME 4. — Quand t tend vers zéro, il existe $0 < \gamma < 1$ tel que $t^\gamma \alpha(t)$ ait une limite stricte positive.

Conséquences. — Une source markovienne vérifie les lois de Zipf (III) et aussi possède la propriété démontrée par Karlin [6] : Z_n^* convenablement normalisée converge faiblement vers une loi de Gauss. Notre modèle, s'il ne décrit que grossièrement la réalité, se trouve donc néanmoins en accord avec certaines tendances décelées depuis longtemps dans les données empiriques.

Démonstration du théorème 4. —

$$\alpha(t) = \sum_{n=1}^{+\infty} \text{card} \{ x \in E^n ; -\log \mu(x) < -\log t \} = \sum_{n=1}^{+\infty} \alpha_n(t) \quad (4.3)$$

or si

$$x \in E^n, \quad n \geq 2, \quad x = x_1 x_2 \dots x_n \quad \mu(x) = p_{\sigma, x_1} p_{x_1, x_2} \dots p_{x_n, \sigma} \quad (4.4)$$

$$-\log \mu(x) = -\log p_{\sigma, x_1} - \log p_{x_n, \sigma} + \sum_{k=1}^{n-1} -\log p_{x_k, x_{k+1}} \quad (4.5)$$

Posons

$$u_k = -\log p_{x_k, x_{k+1}} \quad k = 1, 2, \dots, n-1 \quad (4.6)$$

Les u_k appartiennent à $F = \{ f_{i,j} = -\log p_{i,j} \mid 1 \leq i, j \leq d \}$ F a d^2 éléments ; mais si u_k est donné, u_{k+1} ne peut prendre que d valeurs, celles correspondant aux transitions à partir de x_{k+1} . Si u_1 est donné, il y a exactement d^{n-1} ($n-1$)-uplets possibles (u_2, \dots, u_n) . Nous devons faire du dénombrement sur ces n -uplets. On va donc passer par un calcul sur un modèle

probabiliste sur F assurant à tous les $(n - 1)$ -uplets ci-dessus la même probabilité.

On prend $F^{\mathbb{N}}$, $(U_k)_{k \in \mathbb{N}^*}$ les coordonnées. Si (i, j) désigne $f_{i,j}$ la matrice de transition sera :

$$\Pi = (\pi_{a,b})_{a,b \in F} = \frac{1}{d} \begin{matrix} (1, 1) \dots (1, d) \dots\dots\dots (d, 1) \dots (d, d) \\ \left[\begin{array}{cccccccc} 1 & \dots & 1 & 0 & \dots & 0 & \dots & 0 \\ 0 & \dots & 0 & 1 & \dots & 1 & 0 & \dots & 0 \\ \vdots & & \vdots & & \vdots & & \vdots & & \vdots \\ 0 & \dots & 0 & \dots & 0 & 1 & \dots & 1 & \\ \vdots & & \vdots & & \vdots & & \vdots & & \vdots \\ 1 & \dots & 1 & 0 & \dots & 0 & \dots & 0 & \\ 0 & \dots & 0 & 1 & \dots & 1 & 0 & \dots & 0 \\ \vdots & & \vdots & & \vdots & & \vdots & & \vdots \\ 0 & \dots & 0 & 1 & \dots & 1 & & & \end{array} \right] \begin{matrix} (1, 1) \\ (1, d) \\ (d, 1) \\ (d, d) \end{matrix} \end{matrix} \quad (4.7)$$

c'est-à-dire $\pi_{(i,j),(k,m)} = \frac{1}{d}$ si $j = k$, 0 sinon.

On notera Q^a la probabilité correspondant à l'état initial a , $a \in F$.

$$\alpha_{n+1}(t) = \sum_{\substack{j,k \leq d \\ m,q \leq d}} \text{card} \{ (u_1, \dots, u_n) ; u_1 = f_{j,m}, u_n = f_{q,k} \\ \text{et } u_1 + \dots + u_n < \log(t^{-1} p_{k,\sigma} p_{\sigma,j}) \} \quad (4.8)$$

c'est-à-dire

$$\alpha_{n+1}(t) = \sum_{j,k,m,q \leq d} d^{n-1} Q^{f_{j,m}}(U_n = f_{q,k} \text{ et } U_1 + \dots + U_n < \log(t^{-1} p_{k,\sigma} p_{\sigma,j})) \quad (4.9)$$

Soit $S_n = \sum_1^n U_i$ et si $a, b \in F$

$$G_{a,b}(\theta) = \sum_{k=1}^{+\infty} Q^a(S_k < \theta, X_k = b) d^{k-1} \quad (4.10)$$

alors

$$\sum_{n=1}^{+\infty} \alpha_{n+1}(t) = \sum_{a,b \in F} G_{a,b}(\log(t^{-1} \varphi(a, b))) \quad (4.11)$$

où $\varphi(a, b)$ est défini ainsi : si

$$a = f_{j,m} \quad b = f_{q,k} \quad \varphi(a, b) = p_{k\sigma} p_{\sigma j}. \quad (4.12)$$

Il nous faut donc le comportement asymptotique de $G_{a,b}(\theta)$ quand θ tend vers l'infini. On a les relations

$$\begin{aligned} a \neq b, \quad G_{a,b}(\theta) &= d 1_{\theta > a} \sum_{c \in F} \pi_{a,c} G_{c,b}(t - a) \\ G_{a,a}(\theta) &= 1_{\theta < a} + d 1_{\theta > a} \sum_{c \in F} \pi_{a,c} G_{c,a}(t - a) \end{aligned} \quad (4.13)$$

c'est-à-dire, en passant aux transformées de Laplace :

$$\begin{aligned} \Phi_{a,b}(\lambda) &= \int_0^{+\infty} e^{-\lambda x} G_{a,b}(x) dx \quad \lambda \in \mathbb{C}, \operatorname{Re} \lambda > 0 \\ \Phi(\lambda) &= (\Phi_{a,b}(\lambda))_{a,b \in F}, \quad H(\lambda) = (h_{a,b}(\lambda))_{a,b \in F} \quad h_{a,b}(\lambda) = d \pi_{a,b} e^{-\lambda a} \end{aligned} \quad (4.14)$$

$D(\lambda)$ matrice diagonale $d_a(\lambda) = \frac{e^{-\lambda a}}{\lambda}$, I matrice identité dans F , (4.13) se transforme en

$$\Phi(\lambda)(I - H(\lambda)) = D(\lambda) \quad (4.15)$$

On aperçoit alors une analogie avec des processus de ramification : considérons une population ayant d^2 « types », dont chaque élément produit à la fin de sa vie d enfants, tous de même type, la transition entre le type du géniteur et celui de ses enfants étant régie par la matrice Π (4.7). On suppose que la durée de vie d'un individu ne dépend que de son type. Alors les $G_{a,b}$ interviennent dans le calcul du nombre moyen d'individus en vie à l'instant t (cf. [9] et [1]).

Le résultat final est que :

- 1) $\exists f, g \quad 0 < g < f$ tels que $\forall a, b$

$$G_{a,b}(\theta) = k_{a,b} e^{f\theta} + O(e^{g\theta}) \quad (4.16)$$

où $k_{a,b}$ ne dépend pas de θ

- 2) f est racine de l'équation $\rho(\lambda) = 1$ où $\rho(\lambda)$ est la « plus grande valeur propre (*) de $H(\lambda)$. La fonction ρ est décroissante et continue. (4.17)

Il suffira alors de montrer que $\rho(0) > 1$ et $\rho(1) < 1$ pour pouvoir conclure :

(*) Racine de Frobenius de $H(\lambda)$.

$\alpha(t)$ est somme d'un nombre fixe de termes dont la partie principale, quand t tend vers 0 est en $ke^{f(-\log t)} = kt^f$ avec $0 < f < 1$. f est le γ cherché.

Or soit μ une valeur propre de $H(\lambda)$ et $(w_{i,j})$ un vecteur propre correspondant à μ .

$$H(\lambda)w = \mu w \text{ c'est-à-dire } \sum_{b \in F} h_{a,b}(\lambda)w_b = \mu w_a \text{ ou encore si } a = (i, j)$$

$$\sum_{l=1}^d (p_{i,j})^\lambda w_{j,l} = \mu w_{i,j} \quad (4.18)$$

d'où

$$\mu w_{i,j} = (p_{i,j})^\lambda \sum_l w_{j,l} \text{ et } \sum_{j=1}^d (p_{i,j})^\lambda \left(\sum_l w_{j,l} \right) = \mu \left(\sum_j w_{i,j} \right) \quad (4.19)$$

donc μ est valeur propre de la matrice $C(\lambda) = (c_{i,j}(\lambda))_{i,j \leq d}$ $c_{i,j} = (p_{i,j})^\lambda$ et

$$v = (v_i)_{1 \leq i \leq d} \quad v_i = \sum_{j=1}^d w_{i,j} \quad (4.20)$$

un vecteur propre de $C(\lambda)$ correspondant à μ ; de plus

$$w_{i,j} = \frac{1}{\mu} (p_{i,j})^\lambda v_j \quad \text{si } \mu \neq 0 \quad (4.21)$$

et

$$\forall j \quad \sum_i w_{j,i} = 0$$

définit le sous-espace propre correspondant à la valeur propre 0 de $H(\lambda)$.

Donc 0 est toujours valeur propre de $H(\lambda)$ et les autres sont les valeurs propres non nulles de $C(\lambda)$.

Pour $\lambda = 0$

$$C(0) = \begin{pmatrix} 1 & \dots & 1 \\ \vdots & & \vdots \\ 1 & \dots & 1 \end{pmatrix} \quad (4.22)$$

donc la plus grande valeur propre est d . Donc $\rho(0) > 1$.

Pour $\lambda = 1$

$$C(1) = (p_{i,j})_{1 \leq i,j \leq d} \quad (4.23)$$

or d'après la définition (IV.1) pour tout i , $\sum_{j=1}^d p_{i,j} < 1$ si μ est une valeur

propre positive et v un vecteur propre de $C(1)$ correspondant à μ avec par exemple $|v_1| = \max_{1 \leq j \leq d} |v_j|$

$$\mu |v_1| = |\sum p_{1j} v_j| \leq \sum_1^d p_{1j} |v_j| \leq |v_1| \sum_1^d p_{1j} < |v_1| \quad (4.24)$$

d'où $\rho(1) < 1$.

Ceci termine la démonstration du théorème 4.

V. ÉTUDE DE Z_n^{m*}

Reprenons les notations de II. Il est facile de voir l'équivalence :

$$\{Z_n^{m*} < k\} = \{X_n^{(k)} < m\} \quad (5.1)$$

et donc la proposition 2 entraîne :

COROLLAIRE 5. — $\forall 0 < f < 1$ telle que $\forall k, f \neq p_k$

$$\text{p. s. } Z_n^{(fn)*} \xrightarrow[n \rightarrow +\infty]{} \alpha(f) \quad (5.2)$$

Ceci nous amène à considérer les Z_n^{m*} d'un peu plus près. Karlin [6] a montré que pour m fixé, n tendant vers l'infini $\frac{Z_n^{m*}}{EZ_n^{m*}} \rightarrow 1$ p. s. et a donné un équivalent de EZ_n^{m*} . Le corollaire 5 ci-dessus nous suggère d'étudier le cas n, m tendant simultanément vers l'infini. Cependant si $\frac{m}{n}$ passe par une valeur p_k il y aura problème d'où :

THÉORÈME 6. — Soient l et L , deux réels fixés tels que $0 < l < L < 1$ et $[l, L]$ ne contienne pas de p_k ($k \in \mathbb{N}^*$). Si on pose

$$Z_n^{m*} = \alpha\left(\frac{m}{n}\right) + \Delta_{m,n} \quad (5.3)$$

il existe K et δ positifs tels que

$$\forall m, n \quad l < \frac{m}{n} < L \Rightarrow E(\Delta_{m,n}^2) \leq K e^{-\delta n} \quad (5.4)$$

Démonstration du théorème 6. — Posons

$$A_{s,k} = \{X_n^k = s\} \quad (5.5)$$

$$Z_n^{m*} = \sum_{\substack{s \geq m \\ k \geq 1}} 1_{A_{s,k}}, \quad EZ_n^{m*} = \sum_{\substack{s \geq m \\ k \geq 1}} P(A_{s,k}) \quad (5.6)$$

$$(Z_n^{m*})^2 = Z_n^{m*} + \sum_{\substack{s, t \geq m \\ k, i \geq 1 \\ (s,k) \neq (t,i)}} 1_{A_{s,k}} 1_{A_{t,i}} \quad (5.7)$$

Or si $s \neq t$

$$A_{s,k} A_{t,k} = \emptyset, \quad (Z_n^{m*})^2 - Z_n^{m*} = \sum_{\substack{k \neq i \\ s, t \geq m}} 1_{A_{s,k}} 1_{A_{t,i}} \quad (5.8)$$

d'où

$$E(Z_n^{m*})^2 - EZ_n^{m*} = \sum_{\substack{k \neq i \\ s, t \geq m}} P(A_{s,k} A_{t,i}) \quad (5.9)$$

or

$$P(A_{s,k}) = \binom{n}{s} (p_k)^s (1 - p_k)^{n-s}, \quad s \leq n \quad (5.10)$$

$$P(A_{s,k} A_{t,i}) = \binom{n}{s} \binom{n-s}{i} (p_k)^s (p_i)^t (1 - p_k - p_i)^{n-s-t} \quad (5.11)$$

si $k \neq i$ et $s + t \leq n$.

Il faut dissocier le cas $2m > n$ et $2m \leq n$. On se placera désormais dans le cas $2m \leq n$.

Sur le triangle $0 \leq x \leq 1$, $0 \leq y \leq 1$, $x + y \leq 1$, définissons

$$\Phi(x, y) = \sum_{\substack{s, t \geq m \\ s+t \leq n}} \binom{n}{s} \binom{n-s}{t} x^s y^t (1 - x - y)^{n-s-t} \quad (5.12)$$

alors

$$E(Z_n^{m*})^2 - EZ_n^{m*} = \sum_{k \neq i} \Phi(p_k, p_i) \quad (5.13)$$

si tous les p_k sont inférieurs ou égaux à $\frac{1}{2}$, on a alors

$$E(Z_n^{m*})^2 - EZ_n^{m*} = \sum_{k, i \geq 1} \Phi(p_k, p_i) - \sum_{k \geq 1} \Phi(p_k, p_k) \quad (5.14)$$

De même si on pose

$$\psi(x) = \sum_{s=m}^n \binom{n}{s} x^s (1-x)^{n-s} \quad x \in [0, 1] \quad (5.15)$$

$$EZ_n^{m*} = \sum_{k \geq 1} \psi(p_k) \quad (5.16)$$

Interprétation des fonctions Φ et ψ . — Soient $0 \leq x, y \leq 1$ $x + y \leq 1$

$$I_1 = [0, x], \quad I_2 =]x, 1 - y[, \quad I_3 = [1 - y, 1]$$

et considérons n tirages au sort sur $[0, 1]$, indépendants et suivant la loi uniforme, (U_1, \dots, U_n) et soit $U_n^{(1)} \leq U_n^{(2)} \leq \dots \leq U_n^{(n)}$ le n -uplet ordonné. (5.17)

$$\begin{aligned} \text{Alors } \Phi(x, y) &= \text{P}(\text{au moins } m \text{ points dans } I_1, \text{ au moins } m \text{ points dans } I_3) \\ &= \text{P}(U_n^{(m)} < x, U_n^{(n-m+1)} > 1 - y) \end{aligned} \quad (5.18)$$

$(U_n^{(m)}, 1 - U_n^{(n-m+1)})$ a pour densité (cf. Wilks [10])

$$D_{m,m;n-2m+1}(\xi, \eta) = \frac{n!}{[(m-1)!(n-2m)!]} \xi^{m-1} \eta^{m-1} (1-\xi-\eta)^{n-2m}, \quad 0 < \xi, \eta < 1, \quad \xi + \eta < 1 \quad (5.19)$$

De même $\psi(x) = \text{P}(U_n^{(m)} < x)$ or $U_n^{(m)}$ a pour densité

$$\beta_{m,n-m+1}(\xi) = \frac{n!}{(m-1)!(n-m)!} \xi^{m-1} (1-\xi)^{n-m}, \quad 0 < \xi < 1 \quad (5.20)$$

d'où

$$\begin{aligned} \Phi(x, y) &= \int_0^x \int_0^y D_{m,m;n-2m+1}(\xi, \eta) d\xi d\eta \\ \psi(x) &= \int_0^x \beta_{m,n-m+1}(\xi) d\xi \end{aligned} \quad (5.21)$$

Mais, en intégrant par parties :

$$\begin{aligned} \sum_{k,i} \Phi(p_k, p_i) &= \iint \alpha(dx) \alpha(dy) \Phi(x, y) \\ &= \iint_{\substack{0 \leq x \leq 1 \\ 0 \leq y \leq 1 \\ x+y \leq 1}} \alpha(x) \alpha(y) D_{m,m;n-2m+1}(x, y) dx dy \end{aligned} \quad (5.22)$$

$$\begin{aligned} \sum_k \Phi(p_k, p_k) &= - \int_0^1 \alpha(dx) \Phi(x, x) = \int_0^1 \alpha(x) \left[\frac{\partial \Phi}{\partial x}(x, x) + \frac{\partial \Phi}{\partial y}(x, x) \right] dx \\ &= 2 \int_0^1 \alpha(x) \left[\int_0^x D_{m,m;n-2m+1}(x, \eta) d\eta \right] dx \end{aligned} \quad (5.23)$$

De même

$$\sum_k \psi(p_k) = - \int_0^1 \alpha(dx) \psi(x) = \int_0^1 \alpha(x) \beta_{m, n-m+1}(x) dx \quad (5.24)$$

Or il est bien connu que si G et G' sont deux v. a. indépendantes, de lois gamma de paramètres λ et λ' , $G + G'$ a une loi gamma de paramètre $\lambda + \lambda'$ et $\frac{G}{G + G'}$ une loi bêta de paramètres λ et λ' .

Le résultat analogue à 3 dimensions est le suivant : si G_1, G_2, G_3 sont des v. a. indépendantes, de lois gamma de paramètres $\lambda_1, \lambda_2, \lambda_3$ et si on pose

$$H_1 = \frac{G_1}{G_1 + G_2 + G_3}, \quad H_2 = \frac{G_2}{G_1 + G_2 + G_3},$$

(H_1, H_2) a une loi $D_{\lambda_1, \lambda_2, \lambda_3}$ (et H_1 une loi $\beta_{\lambda_1, \lambda_2 + \lambda_3}$, H_2 une loi $\beta_{\lambda_2, \lambda_1 + \lambda_3}$).

La loi $D_{m, m, n-2m+1}$ (5.19) peut alors être considérée comme la loi d'un couple de v. a. $W_{m, n}, W'_{m, n}$ obtenu de la façon suivante : soient S_m, S'_m, S''_{n-2m+1} 3 v. a. indépendantes, de loi gamma de paramètre m (resp. $m, n - 2m + 1$) et

$$W_{m, n} = \frac{S_m}{S_m + S'_m + S''_{n-2m+1}}, \quad W'_{m, n} = \frac{S'_m}{S_m + S'_m + S''_{n-2m+1}} \quad (5.25)$$

Les formules (5.14) à (5.25) entraînent, si E désigne à la fois l'espérance dans l'espace où ont été définies les v. a. de (5.25) et l'espérance dans l'espace où ont été définies les (Z_n^m) :

$$\begin{aligned} EZ_n^{m*} &= E\alpha(W_{m, n}) \\ E \left[Z_n^{m*} - \alpha\left(\frac{m}{n}\right) \right]^2 &= E \left[\alpha(W_{m, n}) - \alpha\left(\frac{m}{n}\right) \right] \left[\alpha(W'_{m, n}) - \alpha\left(\frac{m}{n}\right) \right] + E\alpha(W_{m, n}) \\ &\quad - 2E[\alpha(W_{m, n}) 1_{W'_{m, n} < W_{m, n}}] \end{aligned} \quad (5.26)$$

L'idée de la suite de la démonstration est la suivante : d'après la loi forte des grands nombres

$$\lim_{m \rightarrow +\infty} \text{p. s.} \frac{S_m}{m} = 1 \quad \lim_{m \rightarrow +\infty} \text{p. s.} \frac{S'_m}{m} = 1 \quad \lim_{n-2m+1 \rightarrow +\infty} \text{p. s.} \frac{S''_{n-2m+1}}{n-2m+1} = 1$$

donc si $\frac{m}{n}$ a une limite f , $\lim_{m, n \rightarrow +\infty} \text{p. s.} (W_{m, n}, W'_{m, n}) = (f, f)$.

α est une fonction en escalier, décroissante, tendant vers l'infini quand son argument tend vers zéro, intégrable :

$$\int_0^1 \alpha(x) dx = \sum_1^{+\infty} p_k = 1$$

Nous allons avoir besoin d'une majoration pour $P\left(\left|W_{m,n} - \frac{m}{n}\right| > \varepsilon\right)$ que nous démontrerons ultérieurement.

LEMME 7. — Soit $W_{m,n}$ une v. a. ayant une loi bêta de paramètres m et $n - m + 1$. Alors :

$$\forall 0 < \varepsilon < 1 \quad \exists N : n \geq N \Rightarrow \forall m \leq \frac{n}{2}$$

$$P\left(\left|W_{m,n} - \frac{m}{n}\right| > \varepsilon\right) \leq 6e^{-\frac{m\varepsilon^2}{128}} \quad (5.27)$$

D'après l'hypothèse du théorème 6, il existe $i \in \mathbb{N}$, $l, L \in [0, 1]$, $N' \in \mathbb{N}$ tels que

$$n, m \geq N' \Rightarrow p_{i+1} < l < \frac{m}{n} < L < p_i \quad (5.28)$$

(on suppose que les p_k sont décroissants pour simplifier ; le cas $l > p_1$ sera vu plus tard). Or

$$\begin{aligned} \alpha(W) &= \alpha\left(\frac{m}{n}\right) & \text{si } p_{i+1} \leq W < p_i \\ \alpha(W) &< \alpha\left(\frac{m}{n}\right) & \text{si } W \geq p_i \\ \alpha(W) &> \alpha\left(\frac{m}{n}\right) & \text{si } W < p_{i+1} \end{aligned} \quad (5.29)$$

(5.26) entraîne donc

$$\begin{aligned} E\left[Z_n^{m*} - \alpha\left(\frac{m}{n}\right)\right]^2 &\leq E\left[\alpha(W_{m,n})\alpha(W'_{m,n})1_{W_{m,n} < p_{i+1}}1_{W'_{m,n} < p_{i+1}}\right] \\ &\quad + \alpha\left(\frac{m}{n}\right)^2 P(W_{m,n} \geq p_i) \\ &\quad + 2\alpha\left(\frac{m}{n}\right)P(W_{m,n} \geq p_i) \\ &\quad + E\left[\alpha(W_{m,n})1_{W_{m,n} < p_{i+1}}\right] \end{aligned} \quad (5.30)$$

Or si $n, m > N'$, $\alpha\left(\frac{m}{n}\right) = i$ fixe

$$P(W_{m,n} \geq p_i) \leq P\left(\left|W_{m,n} - \frac{m}{n}\right| > p_i - \frac{m}{n}\right) \leq P\left(\left|W_{m,n} - \frac{m}{n}\right| > p_i - L\right)$$

donc d'après (5.27)

$$\begin{aligned} n, m \geq \frac{N}{N'} \Rightarrow P(\cdot, W_{m,n} \geq p_i) &\leq 8e^{-\frac{m(p_i-L)^2}{128}} \\ &\leq 8e^{-\frac{l(p_i-L)^2 n}{128}} \end{aligned} \quad (5.31)$$

qui est une majoration du type demandé (5.4).

D'après (5.26) il reste donc à trouver des majorations exponentielles pour

$$\begin{aligned} E_1 &= E\alpha(W_{m,n})\alpha(W'_{m,n})1_{W_{m,n} < p_{i+1}}1_{W'_{m,n} < p_{i+1}} \\ E_2 &= E\alpha(W_{m,n})1_{W_{m,n} < p_{i+1}} \end{aligned} \quad (5.32)$$

$$E_1 = \int_0^{p_{i+1}} \int_0^{p_{i+1}} \alpha(x)\alpha(y)D_{m,m;n-2m+1}(x, y)dx dy$$

soit b quelconque, inférieur à p_{i+1}

$$\begin{aligned} E_1 \leq \int_0^b \int_0^b \alpha(x)\alpha(y)D_{m,m;n-2m+1}(x, y)dx dy + \alpha(b)^2 P(W_{m,n} < p_{i+1}) \\ + 2\alpha(b) \int_0^b \alpha(x)\beta_{m,n-m+1}(x)dx \end{aligned} \quad (5.33)$$

De même

$$E_2 \leq \int_0^b \alpha(x)\beta_{m,n-m+1}(x)dx + \alpha(b)P(W_{m,n} < p_{i+1}) \quad (5.34)$$

On majore $P(W_{m,n} < p_{i+1})$ de la même manière que $P(W \geq p_i)$, à l'aide de (5.27).

Sur $[0, b]^2$, pour $b < \frac{m-1}{n-2}$

$$\begin{aligned} D_{m,m;n-2m+1}(x, y) &\leq \frac{n!}{[(m-1)!]^2(n-2m)!} b^{2m-2}(1-2b)^{n-2m} \\ &= \frac{m^2}{b^2} \underbrace{\frac{n!}{(m!)^2(n-2m)!} b^{2m}(1-2b)^{n-2m}}_{E_3} \end{aligned} \quad (5.35)$$

Et pour $b < \frac{m-1}{n-1}$ $0 < x < b$

$$\beta_{m,n-m+1}(x) < \beta_{m,n-m+1}(b) = \frac{m}{b} \underbrace{\frac{n!}{m!(n-m)!} b^m(1-b)^{n-m}}_{E_4} \quad (5.36)$$

Mais

$$\begin{aligned} \log E_3 &= \log(n!) - 2 \log(m!) - \log(n-2m)! \\ &\quad + 2m \log b + (n-2m) \log(1-2b) \\ \log E_4 &= \log(n!) - \log(m!) - \log[(n-m)!] \\ &\quad + m \log b + (n-m) \log(1-b) \end{aligned} \quad (5.37)$$

or d'après la formule de Stirling

$$\log k! = k \log k - k + \frac{1}{2} \log k + \frac{1}{2} \log 2\pi + \frac{\theta_k}{12k} \quad 0 < \theta_k < 1$$

d'où

$$\begin{aligned} \log E_3 = & -2m \log \frac{m}{n} + 2m \log b - (n-2m) \log \left(1 - \frac{2m}{n}\right) \\ & + (n-2m) \log (1-2b) + \frac{1}{2} \log n - \log m - \frac{1}{2} \log (n-2m) - \log 2\pi \\ & + \frac{1}{12} \left[\frac{\theta_n}{n} - \frac{2\theta_m}{m} - \frac{\theta_{n-2m}}{n-2m} \right] \end{aligned} \quad (5.38)$$

de même

$$\begin{aligned} \log E_4 = & -m \log \frac{m}{n} - (n-m) \log \left(1 - \frac{m}{n}\right) + m \log b + (n-m) \log (1-b) \\ & + \frac{1}{2} \log n - \frac{1}{2} \log m - \frac{1}{2} \log (n-2m) - \frac{1}{2} \log 2\pi \\ & + \frac{1}{12} \left[\frac{\theta_n}{n} - \frac{\theta_m}{m} - \frac{\theta_{n-m}}{n-m} \right] \end{aligned} \quad (5.39)$$

Posant pour $0 < x < \frac{1}{2}$

$$f(x) = -2x \log \frac{x}{b} - (1-2x) \log \frac{1-2x}{1-2b} \quad (5.40)$$

$$g(x) = -x \log \frac{x}{b} - (1-x) \log \frac{1-x}{1-b}$$

on a

$$\begin{aligned} \log E_3 & \leq nf\left(\frac{m}{n}\right) + \frac{1}{2} \log n + \frac{1}{12n} \\ \log E_4 & \leq ng\left(\frac{m}{n}\right) + \frac{1}{2} \log n + \frac{1}{12n} \end{aligned} \quad (5.41)$$

Or f a un maximum égal à 0 pour $x = b$, et après f décroît.

g a un maximum égal à 0 pour $x = b$, et après g décroît.

D'après (5.28), si on prend $b = p_{i+1}$, on a pour $\frac{n}{m} > \frac{N}{N'}$

$$\begin{aligned} \log E_3 & \leq nf(l) + \frac{1}{2} \log n + \frac{1}{12n} \quad \text{avec } f(l) < 0 \\ \log E_4 & \leq ng(l) + \frac{1}{2} \log n + \frac{1}{12n} \quad \text{avec } g(l) < 0 \end{aligned} \quad (5.42)$$

Les formules (5.33) à (5.42), jointes à $\int_0^1 \alpha(x)dx = 1$ permettent alors d'aboutir à l'inégalité (5.4) cherchée.

Ceci dit, nous avons au cours de la démonstration, fait les restrictions $2m \leq n$, $l \leq p_1$ et $p_{(1)} < \frac{1}{2}$. Mais si $2m > n$, $Z_n^{m*} = 0$ ou 1, et donc $EZ_n^{m*} = E(Z_n^{m*})^2$ de même $\alpha\left(\frac{m}{n}\right) = 0$ ou 1.

Alors

$$E\left[Z_n^{m*} - \alpha\left(\frac{m}{n}\right)\right] = \left|E\left(\alpha(W_{m,n}) - \alpha\left(\frac{m}{n}\right)\right)\right|$$

On utilise les mêmes méthodes, on modifie le lemme 7 en changeant m et $n - m + 1$.

On étudie alors les différents cas de figure correspondant à la position respective de

$$p_1, p_2, l, L, \frac{1}{2}.$$

On aboutit à un résultat identique à celui obtenu précédemment.

Démonstration du lemme 7. — On cherche une majoration de $P\left(\left|W_{m,n} - \frac{m}{n}\right| > \varepsilon\right)$. Or $EW_{m,n} = \frac{m}{n+1}$, donc si $n > \frac{1}{\varepsilon}$ et $m \leq \frac{n}{2}$

$$\left|\frac{m}{n+1} - \frac{m}{n}\right| < \frac{\varepsilon}{2} \quad \text{et donc} \quad P\left(\left|W_{m,n} - \frac{m}{n}\right| > \varepsilon\right) \leq P(A_{m,n}) \quad (5.43)$$

où on a posé

$$A_{m,n} = \left\{ \left| W_{m,n} - \frac{m}{n+1} \right| > \frac{\varepsilon}{2} \right\} \quad (5.44)$$

Posons encore, avec les notations (5.25)

$$B_{m,n} = \left\{ S_m - m > -\frac{m}{2} \right\}$$

$$C_{m,n} = \left\{ S'_m + S''_{n-2m+1} - (n-m+1) > -\frac{n-m+1}{2} \right\} \quad (5.45)$$

alors

$$P(A_{m,n}) \leq P(A_{m,n}B_{m,n}C_{m,n}) + P(\bar{B}_{m,n}) + P(\bar{C}_{m,n}) \quad (5.46)$$

Mais

$$A_{m,n} = \left\{ \left| \frac{S_m}{S_m + S'_m + S''_{n-2m+1}} - \frac{m}{n+1} \right| > \frac{\varepsilon}{2} \right\}$$

donc

$$A_{m,n}B_{m,n}C_{m,n} \subset \left\{ \left| S_m \left(1 - \frac{m}{n+1} \right) - (S'_m + S''_{n-2m+1}) \frac{m}{n+1} \right| > \frac{(n+1)\varepsilon}{4} \right\} \quad (5.47)$$

d'où

$$\begin{aligned} P(A_{m,n}B_{m,n}C_{m,n}) &\leq P\left(|S_m - m| > \frac{m\varepsilon}{4}\right) \\ &\quad + P\left(|S'_m + S''_{n-2m+1} - (n - m + 1)| > \frac{(n - m + 1)\varepsilon}{4}\right) \end{aligned} \quad (5.48)$$

car si

$$|S_m - m| \leq \frac{m\varepsilon}{4} \quad \text{et} \quad |S'_m + S''_{n-2m+1} - (n - m + 1)| \leq \frac{n - m + 1}{4} \varepsilon$$

alors

$$\begin{aligned} \left(1 - \frac{m}{n+1}\right) |S_m - m| &\leq \frac{m\varepsilon}{4} \\ \frac{m}{n+1} |S'_m + S''_{n-2m+1} - (n - m + 1)| &\leq \frac{n - m + 1}{4} \varepsilon \end{aligned}$$

et donc

$$\left| S_m \left(1 - \frac{m}{n+1} \right) - (S'_m + S''_{n-2m+1}) \frac{m}{n+1} \right| \leq \frac{m\varepsilon}{4} + \frac{n - m + 1}{4} \varepsilon = \frac{n+1}{4} \varepsilon$$

Nous avons alors besoin d'un lemme sur les v. a. de lois gamma :

LEMME 8. — Si $0 < \sigma < \mu$ et si ξ est une v. a. gamma de paramètre μ alors

$$P(|\xi - \mu| \geq \sigma) \leq 2e^{-\frac{\sigma^2}{8\mu}} \quad (5.49)$$

Démonstration du lemme 8. — Si on pose $M(\varepsilon) = Ee^{\varepsilon(\xi - \mu)}$ défini pour $\varepsilon < 1$, on a

$$M(\varepsilon) = \frac{e^{-\mu\varepsilon}}{(1 - \varepsilon)^\mu} \quad (5.50)$$

si $\varepsilon > 0$, en appliquant l'inégalité de Markov, on a :

$$\forall t > 0 \quad P\left(\xi - \mu > \frac{1}{\varepsilon}(t + \log M(\varepsilon))\right) \leq e^{-t} \quad (5.51)$$

$$P\left(\xi - \mu < -\frac{1}{\varepsilon}(t + \log M(-\varepsilon))\right) \leq e^{-t} \quad (5.51)$$

or si $0 < \varepsilon < \frac{1}{2}$ on a

$$\log M(\varepsilon) \leq \frac{3\mu\varepsilon^2}{2} \quad (5.52)$$

donc

$$\forall t > 0 \quad P\left(\xi - \mu > \frac{1}{\varepsilon} \left(t + \frac{3\mu\varepsilon^2}{2}\right)\right) \leq e^{-t} \quad (5.53)$$

Posant $t = \frac{\sigma^2}{8\mu}$ et $\varepsilon = \frac{\sigma}{2\mu}$ on trouve, si $\sigma < \mu$

$$P(\xi - \mu > \sigma) \leq e^{-\frac{\sigma^2}{8\mu}} \quad (5.54)$$

si $\varepsilon > 0$ on a

$$0 \geq -\log M(-\varepsilon) \geq -\frac{\mu\varepsilon^2}{2} \quad (5.55)$$

donc

$$\forall t > 0 \quad P\left(\xi - \mu \leq -\frac{1}{\varepsilon} \left(t + \frac{\mu\varepsilon^2}{2}\right)\right) \leq e^{-t} \quad (5.56)$$

Posant $t = \frac{\sigma^2}{2\mu}$, $\varepsilon = \frac{\sigma}{\mu}$ il vient

$$P(\xi - \mu < -\sigma) \leq e^{-\frac{\sigma^2}{2\mu}} \quad (5.57)$$

Remarquons que si $\sigma > \mu$ cette dernière probabilité est nulle. Le lemme 8 est donc démontré en réunissant (5.54) et (5.57).

Pour terminer la démonstration du lemme 7, il suffit alors d'appliquer ce qui vient d'être fait à (5.46) et (5.48). En effet :

$$P(\bar{B}_{m,n}) \leq e^{-\frac{m}{8}}, \quad P(\bar{C}_{m,n}) \leq e^{-\frac{n-m+1}{8}} \quad \text{d'après (5.57)}$$

Mais $m < n - m + 1$, d'où

$$P(\bar{B}_{m,n}) + P(\bar{C}_{m,n}) \leq 2e^{-\frac{m}{8}} \quad (5.58)$$

Enfin

$$P\left(|S_m - m| > \frac{m\varepsilon}{4}\right) \leq 2e^{-\frac{m\varepsilon^2}{128}} \quad (5.59)$$

$$P\left(|S'_m + S''_{n-2m+1} - (n-m+1)| > \frac{(n-m+1)\varepsilon}{4}\right) \\ \leq 2e^{-\frac{(n-m+1)\varepsilon^2}{128}} \leq 2e^{-\frac{m\varepsilon^2}{128}} \quad (5.60)$$

(5.58), (5.59) et (5.60) donnent

$$P(A_{m,n}) \leq 6e^{-\frac{m\varepsilon^2}{128}} \quad (5.61)$$

Remarque. — Il ne nous a pas été possible de déduire du théorème 6 des conséquences sur les $X_n^{(k)}$.

2) On peut, avec le théorème 6, estimer la fonction α sur les rationnels et donc, d'une certaine manière, estimer les p_k .

BIBLIOGRAPHIE

- [1] K. B. ATHREYA, P. E. NEY, *Branching processes*, Springer-Verlag, 1972.
- [2] B. BRAINERD, On the relation between types and tokens in a literary text. *J. Appl. Prob.*, t. 9, 1972, p. 507-518.
- [3] B. M. HILL, The rank-frequency form of Zipf's law. *J. Amer. Stat. Assoc.*, vol. 69, no. 348, 1974, p. 1017-1026.
- [4] Hill-Woodrooffe stronger forms of Zipf's law. *J. Amer. Stat. Assoc.*, vol. 70, no. 349, 1975, p. 212-219.
- [5] B. M. HILL, M. WOODROOFE, On Zipf's law. *J. of Appl. Prob.*, t. 12, 1975, p. 425-434.
- [6] S. KARLIN, Central limit theorems for certain infinite urn schemes. *J. of Math. and Mech.*, vol. 17, no 4, 1967, p. 373-401.
- [7] B. MANDELBROT, On the theory of word frequencies and on related markovian models of discourse. *Symposium an Applied Mathem.*, vol. XII, 1962.
- [8] D. N. MCNEIL, Estimating an author's vocabulary. *J. of Amer. Stat. Assoc.*, vol. 68, no. 341, 1973, p. 92-96.
- [9] C. J. MODE, *Multitype branching processes*. American Elsevier, New York, 1971.
- [10] S. S. WILKS, *Mathematical Statistics*. Wiley, New York, London, 1962.

(Manuscrit reçu le 8 décembre 1977)

* Épreuves non corrigées par l'auteur.

Note supplémentaire

Depuis la parution de cet article, Monsieur Benoît Mandelbrot m'a signalé qu'il avait énoncé le résultat du théorème 4, p. 174 sous une forme voisine et donné les grandes lignes de la démonstration lors d'une conférence intitulée « *On recurrent noise limiting coding* » lors du Symposium on Information Networks, Polytechnic Institute of Brocklyn en 1954.

SUPPLEMENT A : LOIS DE ZIPF ET SOURCES MARKOVIENNES

§ I. PRESENTATION.

L'étude précédente mérite une actualisation car au moment où elle a été faite, l'optique grandes déviations m'était inconnue.

Dans la partie IV, le théorème 4 donne une évaluation asymptotique de la queue de la loi de probabilité sur les mots produits par une source markovienne. La preuve que nous avons proposée s'appuie sur une analogie avec les processus de branchements. Si on cherche le fond de l'argument, au-delà du résultat on peut faire 2 remarques.

1) Le caractère markovien de la source donne au logarithme de la probabilité d'un mot (aléatoire) une structure additive (4.5) dont le terme général rentre dans la catégorie des "variables aléatoires définies sur une chaîne de Markov" [1'], [3'] , [4'] .

2) Le mélange de mots de longueur différente dont la probabilité, dans son logarithme, atteint un certain niveau est une situation type renouvellement.

Les calculs analytiques de [3'], [4'] , ainsi que ceux de [1'] (orientés vers le renouvellement), dont j'ai pris connaissance après ce travail et évidemment ceux de Mode [5'] , sont en fait des calculs de grandes déviations. Comme toujours ils contiennent un changement de probabilité recentrant le problème, dont le choix ne devient clair que si l'on se penche sur la question des lois empiriques, telle qu'elle a été résolue par Donsker-Varadhan [2'] (et pour mémoire Sanov [6'] dans le cas indépendant).

La parenté étant ainsi établie, nous sommes en mesure dans § II de proposer pour le théorème 4 une démonstration nouvelle et éclairante, dont une idée se retrouve dans le chap. D . On majore $\alpha(e^{-t})$ par une expression du type $\sum_n \exp - nh \left(\frac{t}{n}\right)$, $h > 0$; pour t grand elle est en $e^{-\gamma t + \sigma(t)}$, où $\gamma = - \inf \frac{h(x)}{x}$ est atteint en β , puis on minore $\alpha(t)$ par $\exp - n_t h \left(\frac{t}{n_t}\right)$ où $n_t = \left[\frac{t}{\beta}\right]$ correspond ipso-facto à une "longueur prépondérante". La preuve sera faite avec une chaîne markovienne d'ordre quelconque, ce qui n'est pas plus difficile.

Dans la partie V, la convergence (exponentielle) de l'estimateur de α a nécessité des calculs intermédiaires : le lemme 8 et sa preuve constituent une redite par rapport à l'inégalité de Chernov ; les évaluations (5.35) et (5.36) pourraient se poursuivre avec des grandes déviations sur des bi-(ou tri-)nomiales.

§ II - A PROPOS DU THEOREME 4.

1. Notations et résultats.

Soit r un entier supérieur ou égal à 2, E un ensemble fini de cardinal $d > 2$, f une fonction de E^r dans \mathbb{R}_+^* , f_* et f^* le minimum et le maximum de f sur E^r . Un élément x de E^n ($n \geq 1$) est noté $x = x_1 \dots x_n$. On définit pour $t > 0$ et $n \geq r$

$$c_n(t) = \text{card} \{x \in E^n ; \sum_{i=1}^n f(x_i, \dots, x_{i+r-1}) < t\}$$

$$c(t) = \sum_{n=r}^{+\infty} c_n(t)$$

Proposition 1 :

$$\lim_{t \rightarrow +\infty} t^{-1} \text{Log } c(t) = \gamma$$

où α est solution d'une équation $L(-\alpha) = 0$, avec L fonction convexe analytique. De plus si $\beta = L'(-\alpha)$ et $n_t = [\frac{t}{\beta}]$ alors on a

$$\lim_{t \rightarrow +\infty} t^{-1} \text{Log } c_{n_t}(t) = \gamma$$

Remarque :

Dans le théorème 4, on a $r = 2$, $f(e_i, e_j) = -\text{Log } p_{e_i, e_j}$ et $c(t)$ représente le nombre de mots ayant une probabilité dépassant e^{-t} . La proposition permet de conclure que $\text{Log } \alpha(x) \underset{x \rightarrow 0}{\sim} \text{Log } (x^{-\gamma})$ (plus faible que 2.2) et, ce qui est nouveau, que parmi les mots de probabilité dépassant x , il y a une plage de longueurs prépondérantes, celles d'ordre équivalent à $\frac{1}{\beta} \text{Log } \frac{1}{x}$.

2. Preuve de la proposition :

Soit $\Omega = E^{\mathbb{N}^*}$, \mathcal{B} la tribu engendrée par les coordonnées $(X_k)_{k \in \mathbb{N}^*}$, ν la loi uniforme sur E , et $P = \nu^{\otimes \mathbb{N}^*}$ sur (Ω, \mathcal{B}, P) la suite $U_k = (X_k, \dots, X_{k+r-1})$ $k = 0, 1, \dots$ est une chaîne de Markov sur E^r de loi initiale $\pi^0 = \nu^{\otimes r}$ et de transition :

$$\begin{aligned} \Pi((e_1, \dots, e_r), (e'_1, \dots, e'_r)) &= d^{-1} \quad \text{si } (e'_1, \dots, e'_{r-1}) = (e_2, \dots, e_r) \\ &= 0 \quad \text{sinon} \end{aligned}$$

Elle est irréductible (Π^r a tous ses éléments strictement positifs) et de loi invariante π^0 . Pour $n \geq 1$, la loi empirique $\xi_n = \frac{1}{n} (\delta_{U_1} + \dots + \delta_{U_n})$ est élément (aléatoire) de $M_1(E^r)$, ensemble des probabilités sur E^r et donc si on note pour $x \in \mathbb{R}_+^*$

$$\Gamma(x) = \{\mu \in M_1(E^r) : \mu(f) < x\}$$

on peut écrire

$$c_{n+p-1}(t) = d^{n+p-1} P(\xi_n \in \Gamma(\frac{t}{n})) \quad (1')$$

La technique suivante s'utilisera aussi au Chap.D. On fixe t ; si n est suffisamment grand, la loi limite ν_0 n'appartient pas à $\Gamma(\frac{t}{n})$ et on a une majoration de type exponentiel pour $P(\xi_n \in \Gamma(\frac{t}{n}))$, pour les autres on majore par 1. On cherche n_t atteignant asymptotiquement la borne (principe " $\Sigma = \sup$ ") et on minore $c(t)$ par $c_{n_t}(t)$.

Procédons au changement de transition, commenté au 3).

Lemme :

Il existe une fonction L_0 de \mathbb{R} dans \mathbb{R} , convexe et analytique et pour tout θ réel une fonction β^θ de E^r dans \mathbb{R} telle que Π^θ définie de $E^r \times E^r$ dans \mathbb{R} par

$$\Pi^\theta(e, e') = \Pi(e, e') \exp(\beta^\theta(e') - \beta^\theta(e) + \theta f(e) - L_0(\theta)) \quad (2')$$

soit une transition, de probabilité invariante π^θ vérifiant

$$\pi^\theta(f) = L'_0(\theta) \quad (3')$$

Preuve du lemme :

Si $\tilde{\Pi}^\theta(e, e') = \Pi(e, e') \exp \theta f(e)$, alors $(\tilde{\Pi}^\theta)^r$ a tous ses éléments strictement positifs et donc d'après [4'], $\tilde{\Pi}^\theta$ possède :

- une plus grande valeur propre (positive) qu'on notera $\exp L_0(\theta)$ avec L_0 convexe et analytique
- un vecteur propre à droite et un vecteur propre à gauche, de coordonnées strictement positives associées à $\exp L_0(\theta)$ (uniques à un facteur près) qu'on notera $\{\exp \beta^\theta(e)\}_{e \in E^r}$ pour le droit et $\{\pi^\theta(e) \exp -\beta^\theta(e)\}_{e \in E^r}$ pour le gauche.

En dérivant les relations définissant les vecteurs propres on obtient (3').

Remarque :

(3') entraîne pour tout $\theta : 0 < f_* \leq L'_0(\theta) \leq f^*$, ce qui, joint à la convexité de L_0 et à la propriété évidente $L_0(0) = 0$ entraîne l'existence d'un unique $\gamma > 0$ vérifiant $L_0(-\gamma) = -\text{Log } d$ (relation type Malthus), ou encore en posant $L = L_0 + \text{Log } d$ et en appelant h la duale de L

$$\gamma = - \inf_x \frac{h(x)}{x}, \quad L(-\gamma) = 0$$

cet infimum est atteint en $\beta = L'(-\gamma) < \pi^0(f)$ (voir dessins)

Suite de la preuve de la proposition :

a) Majoration :

Désignons par P^θ la probabilité sur $(E^r)^{\mathbb{N}^*}$ donnant aux coordonnées U_1, \dots, U_n, \dots une structure de chaîne de Markov de loi initiale π^0 et de transition Π^θ . Si $F_n = \sigma(U_1, \dots, U_n)$ on a toujours

$$\frac{dP^\theta}{dP^\theta} \Big|_{F_n} = \exp \{ \beta^\theta(U_1) - \beta^\theta(U_n) + n L_0(\theta) - n \theta \xi_n(f) \}$$

et donc pour tout $\theta < 0$, $a > 0$

$$P(\xi_n \in \Gamma(a)) \leq \exp \{ n L_0(\theta) - n \theta a + \max_{e, e'} (\beta^\theta(e) - \beta^\theta(e')) \} \quad (4')$$

Or $(\hat{\Pi}^\theta)^r$ admet $\exp r L_0(\theta)$ comme valeur propre avec les mêmes vecteurs propres que $\hat{\Pi}^\theta$. De plus pour $\theta < 0$, chaque terme de $(\hat{\Pi}^\theta)^r$ est minoré par $d^{-p} \exp r \theta f^*$.

On en déduit facilement

$$\max_{e, e'} \{ \beta^\theta(e) - \beta^\theta(e') \} \leq r (L(\theta) - \theta f^*) \quad (5')$$

On obtient donc, grâce à (1'), (4'), (5')

$$c_{n+r-1}(t) < \exp - (n+r) \{ \theta a' - L(\theta) \} \text{ où } a' = \frac{t+r f^*}{n+r}$$

Si $a' \leq \pi^0(f)$ on a $\sup_{\theta \leq 0} \theta a' - L(\theta) = \sup_{\theta} \theta a' - L(\theta) = h(a')$

c'est-à-dire si $N = \inf \{ n : t+r f^* \leq \pi^0(f) (n+r) \}$

$$n \geq N \Rightarrow c_{n+r-1}(t) \leq \exp \gamma(t+r f^*)$$

Et comme pour $t < n f_*$ on a toujours $c_{n+r-1}(t) = 0$, il existe K indépendant de t tel que

$$\sum_{n \geq N} c_{n+r-1}(t) \leq K t e^{\gamma t}$$

La majoration évidente $c_{n+r-1}(t) \leq d^{n+r-1}$ entraîne

$$\sum_{n < N} c_{n+r-1}(t) \leq \frac{1}{d-1} d^{N+r-1}$$

et d'après la définition de N , la dualité, et la définition de α , on a

$$(N+r-1) \text{Log } d < - \frac{t+r f^*}{\pi^0(f)} h(\pi^0(f)) \leq \gamma(t+r f^*)$$

ce qui permet de conclure à l'existence de K' tel que

$$c(t) = \sum_n c_{n+r-1} \leq K' t e^{\gamma t}$$

ou encore $\frac{1}{t} \overline{\text{Lim}} \text{Log } c(t) \leq \gamma$

b) Minoration :

On pose $n_t = \left[\frac{t}{\beta} \right]$ et on minore $c(t)$ par $c_{n_t}(t)$. Or d'après le a) on a pour tout $n \geq 1$, $\lambda \in \mathbb{R}$

$$c_{n+r-1}(t) = d^{n+r-1} E^\theta \frac{1}{\Gamma(\frac{t}{n})} (\xi_n) \exp \{ (\beta^\theta(U_1) - \beta^\theta(U_n)) + n L_0(\theta) - \theta n \xi_n(f) \}$$

Fixons provisoirement $\delta \in]0, \beta - f_*[$ et posons $a = \beta - \frac{\delta}{2}$

$$\lambda = h'(a) < 0 \quad (\text{car } \beta < \pi^0(f)) \quad , \quad \Gamma_\delta = \Gamma(\beta) \setminus \Gamma(\beta - \delta)$$

et $g = \min_{e, e'} \exp \{ \beta^\theta(e) - \beta^\theta(e') \} > 0$

Pour tout $n \leq \frac{t}{\beta}$ on a désormais

$$c_{n+r-1}(t) \geq d^{r-1} g \exp \{ nL(\theta) - \theta n(\beta - \delta) \} P^\theta(\xi_n \in \Gamma_\delta) \quad (6')$$

Le changement de probabilité a été construit de façon à ce que, vu le théorème ergodique on ait

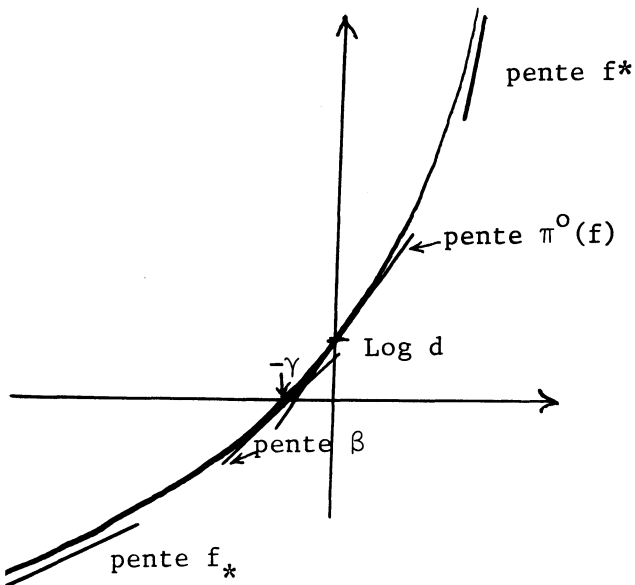
$$\lim_{n \rightarrow +\infty} P^\theta(\xi_n \in \Gamma_\delta) = 1$$

En faisant $n = n_t$ dans (5.3) on obtient donc

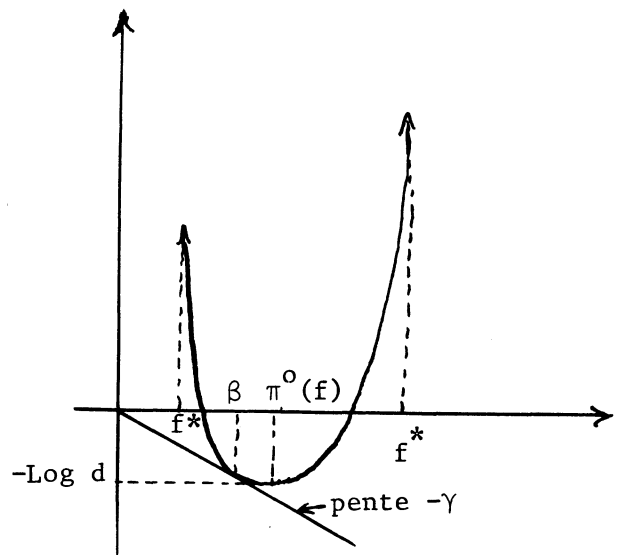
$$\underline{\lim} \frac{1}{t} \text{Log } c(t) \geq \underline{\lim} \frac{1}{t} \text{Log } c_{n_t+r-1}(t) \geq \frac{1}{\beta} [L(\theta) - \theta(\beta - \delta)] \geq - \frac{h(\beta - \delta)}{\beta}$$

et en faisant tendre δ vers 0

$$\underline{\lim} \frac{1}{t} \text{Log } c(t) \geq \gamma$$



Graphe de L .



Graphe de h .

3. Sur le changement de transition :

On se place dans le cadre des "variables aléatoires définies sur une chaîne de Markov finie" ([1'] , [3'] , [4']) ou marches aléatoires typées . Soit F un ensemble fini (c'est le E^r précédent) et $(\mu_{i,j})$ une famille de mesures de Radon positives sur \mathbb{R} , indexée par $F \times F$ et vérifiant

$$\sum_j \mu_{i,j}(1) = 1$$

On fabrique ainsi une transition sur $F \times \mathbb{R}$ puis une chaîne de Markov (R_k, S_k) , $k \in \mathbb{N}$. On note $X_k = S_k - S_{k-1}$ $k \geq 1$, $X_0 = S_0$. Une quantité intéressante est

la loi empirique $L^n = \frac{1}{n} \sum_{k=1}^n \delta_{R_{k-1}, R_k, X_k}$ qui appartient à $M_1(F^2 \times \mathbb{R})$. Si la

matrice $(\mu_{i,j}(1))_{i,j \in F}$ est irréductible et π_0 est sa probabilité invariante, la limite de L^n - théorème ergodique et loi des grands nombres - est évidemment $\pi_0 \otimes \mu$, loi sur $F^2 \times \mathbb{R}$ définie par

$$\pi_0 \otimes \mu (i,j,dx) = \pi_0(i) \mu_{i,j}(dx)$$

A l'aide de [2] on montre l'inégalité de grandes déviations suivantes .

Proposition 2 :

Si B est un borélien de $M_1(F^2 \times \mathbb{R})$ pour la topologie de la convergence étroite on a

$$\forall i \in F \quad \overline{\lim}_n \frac{1}{n} \log P_i(L^n \in B) \leq - I_\mu(\bar{B})$$

où $I_\mu(\bar{B}) = \inf \{K(\pi \otimes \mu, \pi \otimes \nu) ; (\pi \otimes \nu) \in \bar{B} \text{ et } \forall j \sum \pi(i) \nu_{i,j}(1) = \pi(j)\}$

$$\text{et } K(\pi \otimes \mu, \pi \otimes \nu) = \sum_{i,j} \pi(i) \int_{\mathbb{R}} \nu_{i,j}(dx) \log \frac{\nu_{i,j}(dx)}{\mu_{i,j}(dx)} \text{ si } \forall i,j \nu_{i,j} \ll \mu_{i,j}$$

$$= + \infty \quad \text{sinon}$$

A vitesse asymptotique fixée on va retrouver, de façon naturelle (et non plus comme astuce de calcul [3'] , [4']) le changement exponentiel.

Notons pour tout $i, j \in F$, $\theta \in \mathbb{R}$ $\hat{\mu}_{i,j}(\theta) = \int e^{\theta x} \mu_{i,j}(dx)$ supposé fini
 $\rho(\theta)$ la plus grande valeur propre de $M(\theta) = (\hat{\mu}_{i,j}(\theta))_{i,j \in F}$, u^θ et v^θ les
 vecteurs propres à droite et à gauche correspondants, normalisés par
 $\sum_F v^\theta(i) = \sum_F u^\theta(i) v^\theta(i) = 1$. D'après [4'], $L = \text{Log } \rho$ est convexe et analytique,
 soit h sa duale (définie par $h(x) = \sup_\theta \theta x - L(\theta)$). Si a est dans l'inté-
 rieur de l'intervalle de finitude de h alors

$$\inf \{K(\pi \otimes \mu, \pi \otimes \nu) ; (\pi, \nu) \text{ vérifie } \sum_{i,j} \pi(i) \int_{\mathbb{R}} x \nu_{i,j}(dx) = a$$

$$\text{et } \forall j \sum_i \pi(i) \nu_{i,j}(1) = \pi(j)\}$$

est atteint pour

$$\nu_{i,j}(dx) = \frac{u^\theta(j)}{u^\theta(i)} \exp[\theta x - L(\theta)] \mu_{i,j}(dx) \text{ où } \theta = h'(a)$$

Ce résultat qui se prouve à l'aide de multiplicateurs de Lagrange, joint à la
 proposition 2 indique le bon changement de transition pour une étude autour de a .

Bibliographie :

- [1'] ARJAS (E.), SPEED (T.P.), *An extension of Cramer's estimate for the absorption probability of a random walk. Proc. Camb. Phil. Soc., 1973, t.73 , p. 355-359*
- [2'] DONSKER (M.D.), VARADHAN (S.R.S.), *Asymptotic evaluation of certain Markov process expectations for large time. Comm. Pure Appl. Math. 1975, t. 28, p. 1-47 et 1976, t.29, p. 389-461.*
- [3'] KEILSON (J.), WISHART (D.M.G.), *A central limit theorem for processes defined on a finite Markov chain. Proc. Camb. Phil. Soc. 1964, t.60, p. 547-567.*
- [4'] MILLER (H.D.), *A convexity property in the theory of random variables defined on a finite Markov chain. Ann. Math. Stat, 1961, t.32, p. 1260-1270.*
- [5'] MODE (C.J.), *Multitype branching processes. American Elsevier, New-York, 1971.*
- [6'] SANOV (I.N.), *On the probability of large deviations of random variables. Selected Transl. Math. Stat. Prob., 1961, t 1, p. 213-244.*

D. Grandes déviations, dynamique de populations et phénomènes malthusiens.

Grandes déviations, dynamique de populations et phénomènes malthusiens

par

Catherine LAREDO

et

Alain ROUAULT

INRA, Biométrie, Route de Saint-Cyr
78000 Versailles, France

ERA CNRS 532, Univ. Paris-Sud,
Bâtiment 425, 91405 Orsay Cedex, France

ABSTRACT. — An ε -indexed family of Crump-Mode-Jagers branching process with random walk is constructed from a given process, multiplying jumps and life-lengths and dividing frequency of jumps by the same factor ε . Let $\zeta_t^\varepsilon(A)$ be the cardinal of the population lying in A at t . It is shown that either a. s. $\zeta_t^\varepsilon(A) = 0$ for ε small enough or $\varepsilon \text{Log } \zeta_t^\varepsilon(A)$ tends to a positive deterministic limit, in probability as ε tends to 0. That allows to define a time and space homogeneous function α , generalizing the Malthusian parameter of the classical (non spatial) model.

RÉSUMÉ. — Nous considérons une famille, indexée par $\varepsilon > 0$, de processus de branchement de Crump-Mode-Jagers [10] avec marche aléatoire. Le processus correspondant à ε est obtenu à partir d'un processus donné en multipliant par ε les sauts et les durées de vie et en divisant par ε la fréquence des sauts. Soit $\zeta_t^\varepsilon(A)$ l'effectif de la population située dans A à l'instant t . Nous montrons que, ou bien p. s. pour ε assez petit $\zeta_t^\varepsilon(A) = 0$ ou bien $\varepsilon \text{Log } \zeta_t^\varepsilon(A)$ a, en probabilité, une limite déterministe positive quand ε tend vers 0. Ceci permet de définir une fonction homogène α du temps et de l'espace, qui généralise le paramètre de Malthus du modèle classique (non spatial).

I. INTRODUCTION

Un modèle classique de croissance de populations est décrit par un processus de Bellman-Harris [10]. Chaque individu a une durée de vie aléatoire de loi G et à sa mort donne naissance à un nombre aléatoire (d'espérance $m > 1$) d'enfants. Partons d'un seul individu à $t = 0$ et construisons l'arbre généalogique. A l'instant t , des individus de différentes générations sont en vie. Pour t grand, la moyenne empirique des sauts (durées de vie) sur une branche donnée est proche de $\int tG(dt)$ d'après la loi des grands nombres. Les branches correspondant à des moyennes inférieures à $\int tG(dt)$ sont moins probables, mais ont eu le temps de proliférer plus. Il y a phénomène de compensation dont le résultat est qu'asymptotiquement le nombre d'individus en vie à l'instant t est proche de $e^{\alpha t}W$. W est une v. a. positive (strictement sous certaines conditions) et α , appelé *paramètre de Malthus* est l'unique solution positive de

$$(1.1) \quad \int_0^{\infty} e^{-\alpha t} mG(dt) = 1.$$

Ceci se démontre généralement en utilisant la théorie du renouvellement [10]. Nous lui préférons l'approche « grandes déviations » qui peut se résumer ainsi.

Soient X_1, \dots, X_k, \dots des v. a. i. d. de loi G et $S_k = X_1 + \dots + X_k$. Pour tout $\beta > 0$, la contribution de la génération numéro $\left[\frac{t}{\beta} \right]$ ⁽¹⁾ à l'espérance du cardinal de la population en vie à l'instant t est, pour t grand, de l'ordre de

$$\begin{aligned} m^{\left[\frac{t}{\beta} \right]} \mathbf{P}(S_{\left[\frac{t}{\beta} \right]} \simeq t) &= m^{\left[\frac{t}{\beta} \right]} \exp \left(-\frac{t}{\beta} (h_G(\beta) + o(1)) \right) \\ &= \exp \left(-\frac{t}{\beta} (h_G(\beta) - \text{Log } m + o(1)) \right). \end{aligned}$$

(1) $[x]$ désigne le plus grand entier inférieur ou égal à x .

où h_G est la transformée de Cramer [2] de la loi G , c'est-à-dire

$$h_G(\beta) = \sup_{\theta} (\theta\beta - L_G(\theta)) \quad \text{avec} \quad L_G(\theta) = \text{Log} \int_0^{\infty} e^{\theta s} G(ds).$$

Les différentes exponentielles obtenues pour des β différents sont toutes négligeables par rapport à la plus grande $e^{\alpha t}$ où

$$(1.2) \quad \alpha' = - \inf_{\beta > 0} \left(\frac{1}{\beta} (h_G(\beta) - \text{Log } m) \right).$$

Cet inf est atteint en $\beta_0 = \int_0^{\infty} t e^{-\alpha t} m G(dt)$ et $\alpha' = \alpha$. Les générations portant les numéros $\frac{t}{\beta_0} + o(t)$ sont prépondérantes car leurs espérances le sont et une loi des grands nombres joue. β_0 est l'âge moyen de procréation, connu en démographie [10].

Outre cette mise en évidence, cette approche a l'avantage de s'étendre au cas plus général du modèle décrit par Jagers ([10], chap. 6) ⁽²⁾, auquel nous rajoutons une dimension spatiale. A un individu y sont associés : une durée de vie λ_y , un instant de naissance T_y et une position (fixe) X_y . Sa descendance immédiate est décrite par une mesure ponctuelle ξ_y sur $\mathbb{R}^+ \times \mathbb{R}$, donnant les instants de naissance et les positions de ses enfants, calculés à partir de (T_y, X_y) pris comme origine. Nous étudions un processus de branchement où tous les (λ_y, ξ_y) sont i.i.d. Les sauts sur une ligne généalogique prennent leurs valeurs dans $\mathbb{R}^+ \times \mathbb{R}$. L'échelle asymptotique du modèle de Bellman-Harris ($t \rightarrow \infty$) conduit ici à effectuer une homothétie de rapport ε dans $\mathbb{R}^+ \times \mathbb{R}$ et à étudier le comportement de la population quand ε tend vers 0. Autrement dit, à partir d'un processus de référence nous construisons une famille de processus indexée par ε en multipliant par ε les sauts (dans $\mathbb{R}^+ \times \mathbb{R}$) et les durées de vie. Soit $\zeta_t^{\varepsilon}(A)$ le cardinal de la population en vie à l'instant t et située dans A , borélien de \mathbb{R} . Soient μ l'intensité du processus ponctuel ζ , L le logarithme de sa transformée de Laplace, supposée définie partout et h sa transformée de Cramer.

Définissons α , homogène de degré 1, par

$$(1.3) \quad \alpha(t, x) = - \inf_{\beta > 0} \frac{1}{\beta} h\left(\frac{t}{\beta}, \frac{x}{\beta}\right).$$

⁽²⁾ Ce modèle est souvent appelé modèle de Crump-Mode-Jagers.

Nous montrons, sous certaines hypothèses, qu'il existe un cône \mathcal{S} tel que, pour tout intervalle I de \mathbb{R} :

- 1) si $\{t\} \times I$ est inclus dans $(\overline{\mathcal{S}})^c$, alors p. s. $\zeta_\varepsilon^t(I) = 0$ pour ε assez petit;
- 2) si $\{t\} \times I$ coupe \mathcal{S} , alors $\varepsilon \text{Log } \zeta_\varepsilon^t(I)$ converge en probabilité vers $\sup_{x \in I} \alpha(t, x)$ quand ε tend vers 0.

La fonction α joue le rôle de paramètre de Malthus local, et la relation (1.1) devient

$$(1.4) \quad L\left(-\frac{\partial \alpha}{\partial t}, -\frac{\partial \alpha}{\partial x}\right) = 0.$$

Le travail décisif sur l'utilisation des grandes déviations dans les processus de branchement est dû à Biggins [4] [5] [6], dont nous utilisons plusieurs techniques. Après la rédaction de ce travail, nous avons appris [8] (voir aussi [7]) que Biggins avait obtenu un résultat voisin (convergence p. s. mais avec durée de vie infinie). Son optique est un peu différente, il fait tendre t vers l'infini, sans introduire de ε . L'intérêt de ε est d'envisager $\alpha(t, \cdot)$ dans son évolution au cours du temps. Dans le cas présent, il est possible de ne pas introduire de ε car la fonction α est homogène à cause de l'hypothèse d'équidistribution des ζ_y (la reproduction ne dépend ni de l'instant, ni de la position). Cependant ce cadre de travail est adéquat pour étudier le cas général d'une dépendance de ζ_y par rapport à (T_y, X_y) .

II. NOTATIONS ET RÉSULTATS PRÉLIMINAIRES

1. Processus de branchement.

Les notations de base sont proches de Jagers ([10], chap. 6). Nous rajoutons seulement une dimension spatiale sous la forme de l'analogie d'une marche aléatoire sur \mathbb{R} . A un individu né à l'instant σ , situé en a , on associe un couple de v. a. (λ, ξ) où λ est à valeurs dans \mathbb{R}^+ et ξ à valeurs dans M , ensemble des mesures ponctuelles sur $\mathbb{R}^+ \times \mathbb{R}$ muni de la tribu borélienne pour la topologie vague. λ représente la durée de vie de l'individu et ξ son processus de reproduction. Les points de ξ représentent les couples (instant de naissance, position) des différents enfants, à partir de (σ, a) pris pour origine. Nous supposons que la loi de (λ, ξ) , définie sur $\mathbb{R}^+ \times M$ est indé-

pendante de (σ, a) (processus homogène) et que ξ a une intensité μ , mesure de Radon sur $\mathbb{R}^+ \times \mathbb{R}$ telle que

$$\mu(B) = E\xi(B) \quad (B \text{ borélien borné})$$

$$\int_{\mathbb{R}^+ \times \mathbb{R}} f(t, x) \mu(dt, dx) = E \int_{\mathbb{R}^+ \times \mathbb{R}} f(t, x) \xi(dt, dx)$$

(f borélienne bornée).

Nous supposons

$$(H1) \quad 1 < m = \mu(\mathbb{R}^+ \times \mathbb{R}) < \infty$$

(m est le nombre moyen d'enfants par individu)

$$(H2) \quad P(\xi([0, \lambda] \times \mathbb{R}) \geq 1) = 1$$

(un individu a au moins un enfant durant sa vie).

Soient $\mathbb{N}^* = \mathbb{N} \setminus \{0\}$, $J_n = \mathbb{N}^{*n}$, $J = \{0\} \cup \left(\bigcup_{n=1}^{\infty} J_n \right)$. J est l'ensemble potentiel des descendants de $\{0\}$ y compris ce dernier, décomposé en générations. Notons y l'élément général, dont les descendants potentiels sont les z de $J(y) = \{y\} \times J$ (par convention $(y, 0) = y$).

Soit $(\Omega, \mathcal{A}, P) = \prod_{y \in J} (\Omega_y, \mathcal{A}_y, P_y)$ où les $(\Omega_y, \mathcal{A}_y, P_y)$ sont des espaces

identiques sur lesquels on définit (λ_y, ξ_y) de même loi que (λ, ξ) . La notion de réalisation est définie par récurrence:

0 est réalisé;

(y, k) est réalisé si y l'est et $\xi_y(\mathbb{R}^+ \times \mathbb{R}) \geq k$ ($k \in \mathbb{N}^*$).

Il est facile de construire une indexation mesurable dans M , c'est-à-dire une application mesurable de $M \times \mathbb{N}^*$ dans $\mathbb{R}^+ \times \mathbb{R}$ (cf. [12], p. 247) permettant de définir sans ambiguïté « le $k^{\text{ième}}$ point de ξ_y » noté (T_y^k, X_y^k) , pour $k \leq \xi_y(\mathbb{R}^+ \times \mathbb{R})$. L'instant de naissance T_y et la position X_y d'un individu $y = (i_1, \dots, i_n)$ sont obtenus en additionnant les quantités relatives à chacun de ses ancêtres par rapport au précédent :

$$(2.1.1) \quad T_y = T_0^{i_1} + \dots + T_{(i_1, \dots, i_{n-1})}^{i_n}, \quad X_y = X_0^{i_1} + \dots + X_{(i_1, \dots, i_{n-1})}^{i_n}$$

(et $(T_0, X_0) = (0, 0)$).

L'ensemble aléatoire des individus réalisés de J_n (resp. $J_n(y)$) est noté R_n (resp. $R_n(y)$). Le processus obtenu en oubliant les positions, $(\text{card } R_n)_{n \in \mathbb{N}}$ est un processus de Galton-Watson d'espérance m , supercritique, avec non-extinction p. s. sous les hypothèses (H1) et (H2).

Soit $A \subset \mathbb{R}^+ \times \mathbb{R}$; $\eta(A)$ est l'ensemble des « nés dans A », décomposé suivant les générations en les $\eta_n(A)$; $Z(A)$ et $Z_n(A)$ sont les cardinaux de ces ensembles

$$(2.1.2) \quad \eta_n(A) = \{ y \in \mathbb{R}_n \mid (T_y, X_y) \in A \}, \quad \eta(A) = \bigcup_{n=0}^{\infty} \eta_n(A)$$

$$Z_n(A) = \sum_{y \in \mathbb{R}_n} \chi((T_y, X_y) \in A) \quad (3), \quad Z(A) = \sum_{n=0}^{\infty} Z_n(A)$$

Si $t \in \mathbb{R}^+$, $I \subset \mathbb{R}$, le nombre d'individus en vie à l'instant t et situés dans I est $\zeta_t(I)$:

$$(2.1.3) \quad \zeta_t(I) = \sum_{n=0}^{+\infty} \zeta_{t,n}(I); \quad \zeta_{t,n}(I) = \sum_{y \in \mathbb{R}_n} \chi(T_y \leq t < T_y + \lambda_y) \chi(X_y \in I).$$

De même, le nombre de descendants de y en vie à l'instant t et situés dans I est

$$(2.1.4) \quad \zeta_t(I; y) = \sum_{n=0}^{+\infty} \sum_{z \in \mathbb{R}_n(y)} \chi(T_z \leq t < T_z + \lambda_z) \chi(X_z \in I).$$

L'hypothèse H2 peut alors se réécrire

$$(2.1.5) \quad \chi(T_y \leq t) \leq \zeta_t(\mathbb{R}; y) \quad (\text{pour tout } y \text{ et tout } t).$$

Par construction, deux individus différents de la même génération ont des descendance distinctes (principe d'exclusion):

$$(2.1.6) \quad \sum_{y \in \mathbb{R}_n} \zeta_t(I; y) \leq \zeta_t(I) \quad (\text{pour tout } n, y, t).$$

On peut relier les v. a. ζ et Z de la façon suivante. I_1 et I_2 sont 2 intervalles de \mathbb{R} , $I_1 \subset I_2$, u et t vérifient $0 < u < t$ et A est le rectangle $[t-u, t] \times I_1$. Alors, sous (H2), on a :

$$(2.1.7) \quad \zeta_t(I_2) \geq Z_n(A) - \sum_{y \in \eta_n(A)} \zeta_t(I_2^c; y).$$

En effet si $y \in \eta_n(A)$, on a $T_y \leq t$, donc d'après (2.1.5)

$$1 \leq \zeta_t(\mathbb{R}; y) = \zeta_t(I_2; y) + \zeta_t(I_2^c; y).$$

(3) $\chi(\cdot)$ est la fonction indicatrice d'événement.

Il suffit d'appliquer le principe d'exclusion (2.1.6) pour aboutir à (2.1.7).
La famille indexée par ε est définie par

$$(2.1.8) \quad \zeta_{t,n}^\varepsilon(\mathbf{I}) = \zeta_{t,n} \left(\frac{1}{\varepsilon} \mathbf{I} \right), \quad \zeta_t^\varepsilon(\mathbf{I}) = \zeta_t \left(\frac{1}{\varepsilon} \mathbf{I} \right).$$

En combinant (2.1.1) et (2.1.2) on vérifie facilement la relation suivante, base d'une analogie fructueuse avec les marches aléatoires [4]

$$(2.1.9) \quad \text{EZ}_n(\mathbf{A}) = \mu_n(\mathbf{A}) \quad (\mu_n \text{ n}^{\text{ième}} \text{ puissance de convolution de } \mu).$$

Enfin \mathcal{B}_n désigne la tribu engendrée par les n premières générations non comptés leurs temps de vie, c'est-à-dire engendrée par les $\zeta_y, y \in \bigcup_{k=1}^{n-1} \mathbf{J}_k$.

2. Transformée de Cramer.

Considérons une mesure μ positive sur $\mathbb{R}^+ \times \mathbb{R}$, finie. Si $y = (t, x)$ est le point courant de $\mathbb{R}^+ \times \mathbb{R}$, supposons

$$(H3) \quad 1) \text{ pour tout } \theta \text{ de } \mathbb{R}^2 \text{ exp } L(\theta) = \int_{\mathbb{R}^+ \times \mathbb{R}} \exp((\theta, y)) \mu(dy) < \infty (*)$$

2) C étant l'enveloppe convexe fermée du support de μ , on a $\mu(\partial C) = 0$ (donc $C \neq \emptyset$) (*).

$$(2.2.1) \quad h(y) = \sup_{\theta \in \mathbb{R}^2} ((\theta, y) - L(\theta))$$

L est convexe et analytique. h est s. c. i. partout, finie sur $\overset{\circ}{C}$, infinie sur ∂C et C^c , strictement convexe et analytique sur $\overset{\circ}{C}$. $\text{grad } L$ est un difféomorphisme de \mathbb{R}^2 sur $\overset{\circ}{C}$. De plus, pour tout y de $\overset{\circ}{C}$, on a :

$$(2.2.2) \quad h(y) = (\theta, y) - L(\theta)$$

où θ est l'unique solution de $\text{grad } L(\theta) = y$, et alors $\text{grad } h(y) = \theta$. Nous dirons qu'un tel couple (θ, y) est en dualité [2].

Pour tout B sous-ensemble de $\mathbb{R}^+ \times \mathbb{R}$, posons

$$(2.2.3) \quad H(B) = \inf_{y \in B} h(y).$$

(*) (θ, y) désigne le produit scalaire de θ et y dans \mathbb{R}^2 ; $\partial C, \overset{\circ}{C}, \overline{C}, C^c$ sont resp. la frontière, l'intérieur, la fermeture et le complémentaire de C .

Au voisinage de l'axe $\{0\} \times \mathbb{R}$, h a les 2 propriétés suivantes :

1) Si $a \in \mathbb{R}^+$ et si Γ_a^+ et Γ_a^- sont les 2 cônes $\{(t, x) \in \mathbb{R}^+ \times \mathbb{R} \mid x \geq at\}$ et $\{(t, x) \in \mathbb{R}^+ \times \mathbb{R} \mid x \leq -at\}$ alors $\underline{H}(a) = H(\Gamma_a^+ \cup \Gamma_a^-)$ vérifie

$$(2.2.4) \quad \lim_{a \rightarrow +\infty} \underline{H}(a) = +\infty.$$

En effet l'ensemble $K_M = \{y \mid h(y) \leq M\}$ est compact car convexe fermé borné dans toutes les directions (Banach-Steinhaus) et ne touche pas $\{0\} \times \mathbb{R} \subset (\overset{\circ}{C})^c$. Il existe donc $a > 0$ tel que $K_M \cap (\Gamma_a^+ \cup \Gamma_a^-) = \emptyset$ (voir fig. 1).

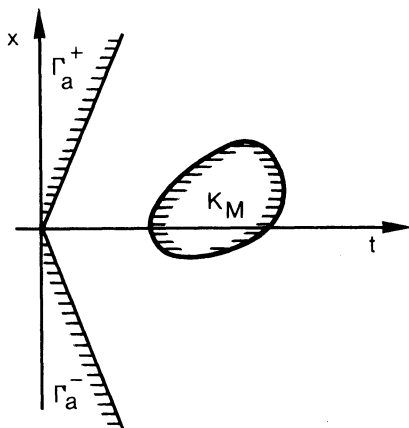


FIG. 1.

2) Si B est compact et $k > 0$, alors

$$(2.2.5) \quad \lim_{k \rightarrow +\infty} H\left(\frac{1}{k} B\right) = +\infty.$$

Il suffit en effet de remarquer qu'il existe k tel que $\frac{1}{k} B \cap K_M = \emptyset$.

Le minimum de h est obtenu pour $\theta = 0$, il vaut $-L(0) = -\text{Log } m$. Le cône $\mathcal{S} = \{y \mid \text{il existe } \beta > 0 \text{ tel que } h(\beta y) < 0\}$ est donc non vide sous (H1) et la fonction α définie par

$$(2.2.6) \quad \alpha(y) = \sup_{\beta > 0} \left(-\frac{1}{\beta} h(\beta y) \right)$$

est finie sur $\mathcal{C} = \mathbb{R}^+ * \overset{\circ}{C}$ et strictement positive sur \mathcal{S} , négative sur \mathcal{S}^c .

THÉORÈME 1. — Sous l'hypothèse (H3), il existe une fonction β définie

sur \mathcal{C} , à valeurs positives telle que $\beta(y)\alpha(y) = -h(\beta(y)y)$. Sur \mathcal{C} , α et β sont de classe \mathcal{C}^∞ , β est homogène de degré -1 , α est homogène de degré 1 et vérifie

$$(2.2.7) \quad L(-\text{grad } \alpha) = 0$$

Démonstration. — Sur \mathcal{C} , h est finie, de sorte que $\frac{1}{\beta}h(\beta y)$ a un minimum à distance finie, comme l'indique le dessin, h étant convexe et infinie aux deux extrémités de $\mathbb{R}^+ \cdot y$; la tangente réalise la pente minimum, l'abscisse est $\beta \cdot y$; les homogénéités sont bien claires. Au minimum, on a successivement $h(\beta y) = (\beta y, \text{grad } h(\beta y))$ d'où

$$L(\text{grad } h(\beta y)) = 0 \quad \text{et} \quad -\alpha\beta = h(\beta y) = (\beta y, \text{grad } h(\beta y))$$

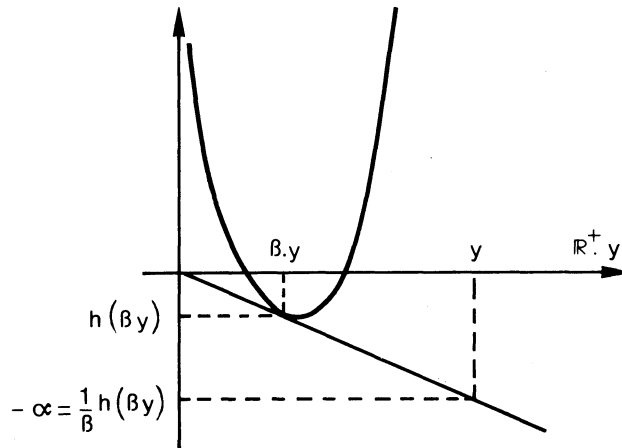


FIG. 2.

d'où $-\alpha = (y, \text{grad } h(\beta y))$ ce qui entraîne $\text{grad } \alpha(y) = -\text{grad } h(\beta y)$ par la formule d'Euler et permet de conclure à la validité de (2.2.7).

3. Encadrements à la Cramer-Chernov.

a) *Changement de loi :*

Soient θ et y en dualité suivant (2.2.2). Posons

$$\mu^\theta(dz) = \exp((\theta, z) - L(\theta))\mu(dz)$$

C'est une probabilité, de transformée de Laplace analytique valant en s : $\exp(L(\theta + s) - L(\theta))$. Le moment du 1^{er} ordre s'obtient par dérivation et vaut $\text{grad } L = y$. La variance est finie, continue par rapport à y (ou θ) (c'est en effet le Hessien de L) et de rang plein (sinon une combinaison affine est p. p. nulle et on est en dimension inférieure, ce qui contredit (H3) 2)). On appelle $\bar{\mu}^\theta$ la probabilité centrée d'expression :

$$(2.3.1) \quad \begin{aligned} \bar{\mu}^\theta(dz) &= \exp((\theta, y + z) - L(\theta)) \cdot \mu(y + dz) \\ &= \exp(h(y)) \cdot \exp(\theta, z) \cdot \mu(y + dz). \end{aligned}$$

Il est immédiat de vérifier que sa $n^{\text{ième}}$ puissance de convolution est

$$(2.3.2) \quad \begin{aligned} (\bar{\mu}^\theta)_n(dz) &= \exp((\theta, ny + z) - nL(\theta)) \cdot \mu_n(ny + dz) \\ &= \exp(nh(y)) \cdot \exp(\theta, z) \cdot \mu_n(ny + dz). \end{aligned}$$

C'est aussi $(\bar{\mu}_n)^\theta(dz)$, car θ et ny sont duales pour μ_n . Nous la noterons définitivement $\bar{\mu}_n^\theta$.

D'après (2.1.2), Z_n est un élément aléatoire de M ; la transformation de μ_n en $\bar{\mu}_n^\theta$ conduit à définir :

$$(2.3.3) \quad \begin{aligned} \bar{Z}_n^\theta(dz) &= \exp((\theta, ny + z) - nL(\theta)) \cdot Z_n(ny + dz) \\ &= \exp(nh(y)) \cdot \exp(\theta, z) \cdot Z_n(ny + dz). \end{aligned}$$

Les propriétés suivantes sont immédiates (A est un borélien de $\mathbb{R}^+ \times \mathbb{R}$)

$$(2.3.4) \quad Z_n(A) = \exp(-nh(y)) \cdot \int_{A - ny} \exp(-(\theta, z)) \cdot \bar{Z}_n^\theta(dz)$$

$$(2.3.5) \quad E\bar{Z}_n^\theta(A) = \bar{\mu}_n^\theta(A) \quad (\text{analogue de 2.1.9})$$

b) *Inégalités* :

Soit B un convexe borné, inclus dans $\overset{\circ}{C}$. Le sup de h sur B est atteint en y de ∂B , de dual θ . Si $u \in B$, μ p. p. on a : $(\theta, u - y) \leq 0$ (voir fig. 3). (2.3.3) entraîne pour tout B :

$$(2.3.6) \quad Z_n(nB) \geq \exp(-nh(y)) \cdot \bar{Z}_n^\theta(n(B - y))$$

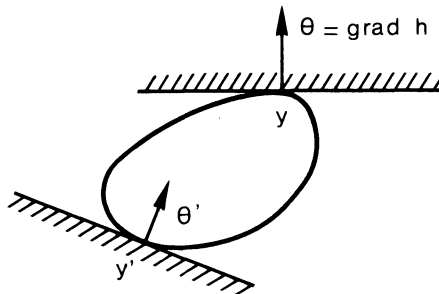


FIG. 3.

et par intégration (cf. (2.1.9))

$$(2.3.6') \quad \mu_n(n\mathbf{B}) \geq \exp(-nh(y)) \cdot \bar{\mu}_n^\theta(n(\mathbf{B}-y)).$$

Si \mathbf{B} ne contient pas le minimum de h , l'inf de h sur \mathbf{B} est atteint en y' sur $\partial\mathbf{B}$, de dual θ' . Si $u \in \mathbf{B}$, on a μ p. p. $(\theta', u-y) \geq 0$ d'où

$$(2.3.7) \quad Z_n(n\mathbf{B}) \leq \exp(-nh(y')) \cdot \bar{Z}_n^{\theta'}(n(\mathbf{B}-y'))$$

$$(2.3.7') \quad \mu_n(n\mathbf{B}) \leq \exp(-nh(y')) \cdot \bar{\mu}_n^{\theta'}(n(\mathbf{B}-y')).$$

Enfin si \mathbf{B} contient le minimum de h (égal à $-\text{Log } m$), on a évidemment $u_n(n\mathbf{B}) \leq \mu_n(\mathbb{R}^+ \times \mathbb{R}) = m^n$ donc dans tous les cas :

$$(2.3.8) \quad \mu_n(n\mathbf{B}) \leq \exp(-nH(\mathbf{B}))$$

Lorsque μ est une probabilité (2.3.6'), (2.3.7') et (2.3.8) sont des formules de base de la théorie des grandes déviations. Enfin nous utiliserons, en vue de l'étude de domaines variables, un résultat voisin de (2.3.6).

Soit $y \in \mathring{\mathbf{C}}$, $\theta = (\theta_1, \theta_2)$ son dual, $|\theta| = |\theta_1| + |\theta_2|$, $\mathbf{B} = \mathbf{I} \times \mathbf{J}$ un rectangle contenant y et $\rho = \max(|\mathbf{I}|, |\mathbf{J}|)$. On a

$$(2.3.9) \quad Z_n(n\mathbf{B}) \geq \exp(-n(h(y) + \rho|\theta|)) \cdot \bar{Z}_n^\theta(n(\mathbf{B}-y)).$$

Il suffit en effet dans (2.3.4) de constater que $\mathbf{B}-y$ contient 0 et donc que $|(\theta, z)| \leq |\theta| n\rho$.

III. RÉSULTAT PRINCIPAL

Énonçons encore 2 hypothèses :

(H4) La loi G de la v. a. λ (durée de vie) a une transformée de Laplace finie partout $\left(\hat{G}(\eta) = \int_0^\infty e^{-\eta t} G(dt) < \infty \text{ pour tout } \eta \right)$.

(H5) Il existe $k > 1$, tel que, si on pose $g_k(x) = x[(\text{Log}^+ x)^k + 1]$ pour $x \in \mathbb{R}^+$, on ait pour tout $\theta \in g_k(\bar{Z}_1^\theta(\mathbb{R}^+ \times \mathbb{R})) < \infty$.

THÉORÈME 2. — Soit \mathbf{I} un intervalle fermé de \mathbb{R} .

a) (Sous les hypothèses H3 et H4). Si $\{t\} \times \mathbf{I}$ est inclus dans $(\mathcal{P})^c$ alors p. s. il existe ε_0 tel que pour $0 < \varepsilon < \varepsilon_0$ on ait $\zeta_\varepsilon^i(\mathbf{I}) = 0$.

b) (Sous les hypothèses H1 à H5). Si $\{t\} \times \mathbf{I}$ a une intersection non vide avec \mathcal{S} , $\varepsilon \text{Log } \zeta_\varepsilon^i(\mathbf{I})$ converge en probabilité vers $\sup_{x \in \mathbf{I}} \alpha(t, x)$ quand ε tend vers 0.

Comme dans tous les processus de branchement il y a alternative entre

extinction et croissance exponentielle (sur \mathcal{S} , α est positive). La difficulté réside ici dans la présence simultanée de plusieurs générations. La décomposition (2.1.3) et les majorations de Cramer-Chernov mettent en évidence un majorant exponentiel pour $E\zeta_t^\varepsilon(\mathbf{I})$ dont la contribution essentielle est apportée par les n de l'ordre de $\frac{C}{\varepsilon} + o\left(\frac{1}{\varepsilon}\right)$ pour un c convenable. La démonstration de *b*) nécessite une minoration obtenue en isolant un tel n dans (2.13), en comparant $\zeta_t^\varepsilon(\mathbf{I})$ à un $Z_n(nB_n)$ pour B_n convenable grâce à (2.1.7), en extrayant la partie exponentielle grâce à (2.3.9) et en appliquant le principe de la génération intermédiaire (cf. [6] où est faite une étude précise, dans \mathbb{R} de $\int g(x - nb)Z_n(dx)$ pour b et g fixés).

La démonstration du théorème 2 utilise le lemme suivant.

LEMME 1. — Sous les hypothèses (H3) et (H4), si \mathbf{I} est un intervalle fermé tel que $\{t\} \times \mathbf{I}$ ait une intersection non vide avec \mathcal{S} on a

$$(3.1) \quad \overline{\lim}_{\varepsilon \rightarrow 0} \varepsilon \operatorname{Log} E\zeta_t^\varepsilon(\mathbf{I}) \leq \sup_{x \in \mathbf{I}} \alpha(t, x).$$

Démonstration du lemme. — Si pour $0 < \tau < t$ on note $B_\tau = [t - \tau, t] \times \mathbf{I}$ et $B_t = [0, t] \times \mathbf{I}$ on a d'après (2.1.3) pour tout n, t, ε

$$(3.2) \quad \zeta_{t,n}^\varepsilon(\mathbf{I}) \leq Z_n\left(\frac{1}{\varepsilon} B_t\right), \quad \zeta_{t,n}^\varepsilon(\mathbf{I}) \leq Z_n\left(\frac{1}{\varepsilon} B_\tau\right) + \sum_{y \in \mathbb{R}_n} \chi\left(\lambda_y > \frac{\tau}{\varepsilon}\right)$$

d'où pour $N > 0$ quelconque, d'après (2.1.8)

$$(3.3) \quad E\zeta_t^\varepsilon(\mathbf{I}) \leq \sum_{n=1}^N \mu_n\left(\frac{B_\tau}{\varepsilon}\right) + \sum_{n=0}^N E \sum_{y \in \mathbb{R}_n} \chi\left(\lambda_y > \frac{\tau}{\varepsilon}\right) + \sum_{n=N+1}^{+\infty} \mu_n\left(\frac{1}{\varepsilon} B_t\right).$$

Or les $\lambda_y, y \in J_n$, sont indépendants de \mathcal{B}_n et de loi G , et les $\chi(y \in \mathbb{R}_n)$ pour $y \in J_n$ sont \mathcal{B}_n mesurables d'où pour tout $n \geq 0$ et tout $d > 0$

$$(3.4) \quad E \sum_{y \in \mathbb{R}_n} \chi\left(\lambda_y > \frac{\tau}{\varepsilon}\right) = m^n G\left(\left[\frac{\tau}{\varepsilon}, +\infty\right]\right) \leq m^n \hat{G}(d) \exp\left(-\frac{d\tau}{\varepsilon}\right)$$

grâce à la majoration exponentielle classique et (H4). Appliquant (2.3.8) à chacun des μ_n il vient

$$(3.5) \quad E\zeta_r^\varepsilon(\mathbf{I}) \leq \sum_{n=1}^N \exp\left(-nH\left(\frac{1}{n\varepsilon} \mathbf{B}_\tau\right)\right) + \frac{m^{N+1}}{m-1} \hat{G}(d) \exp\left(-\frac{d\tau}{\varepsilon}\right) \\ + \sum_{n=N+1}^{+\infty} \exp\left(-nH\left(\frac{1}{n\varepsilon} \mathbf{B}_\tau\right)\right).$$

D'après (2.2.5), pour tout $r > 0$ on peut trouver k_0 tel que $H\left(\frac{1}{k} \mathbf{B}_\tau\right) \geq r$ pour $k \geq k_0$. Le choix de $N = \left[\frac{k_0}{\varepsilon}\right]$ conduit à :

$$(3.6) \quad \sum_{n=N+1}^{+\infty} \exp\left(-nH\left(\frac{1}{n\varepsilon} \mathbf{B}_\tau\right)\right) \leq \sum_{n=N+1}^{+\infty} e^{-nr} = O(e^{-Nr}).$$

Pour majorer $\sum_{n=1}^N \exp\left(-nH\left(\frac{1}{n\varepsilon} \mathbf{B}_\tau\right)\right)$ on écrit :

$$(3.7) \quad nH\left(\frac{1}{n\varepsilon} \mathbf{B}_\tau\right) \geq \frac{1}{\varepsilon} \inf_{n>0} \left(n\varepsilon \inf_{y \in \mathbf{B}_\tau} h\left(\frac{y}{n\varepsilon}\right)\right) \geq \frac{1}{\varepsilon} \inf_{y \in \mathbf{B}_\tau} (-\alpha(y))$$

d'après la définition de α (2.2.6). En regroupant on obtient

$$(3.8) \quad E\zeta_r^\varepsilon(\mathbf{I}) \leq \frac{k_0}{\varepsilon} \exp\left(\frac{1}{\varepsilon} \sup_{y \in \mathbf{B}_\tau} \alpha(y)\right) + \frac{1}{m-1} \hat{G}(d) \exp\left(\frac{1}{\varepsilon} (k_0 \text{Log } m - d\tau)\right) \\ + O(e^{-Nr}).$$

Puisque la fonction α vaut $-\infty$ à l'extérieur de \mathcal{C} , est négative sur \mathcal{S}^c et positive sur \mathcal{S} et que $\mathbf{B}_0 \cap \mathcal{S} \neq \emptyset$, le sup de α sur \mathbf{B}_τ est le même que sur $\mathbf{B}_\tau \cap \mathcal{S}$ donc strictement positif, et εLog du premier terme de (3.8) tend vers $\sup_{y \in \mathbf{B}_\tau \cap \mathcal{S}} \alpha(y) > 0$. Pour tout $\tau > 0$, les deux autres termes sont bornés pour le choix $d > \frac{k_0 \text{Log } m}{\tau}$ (possible d'après (H4)). Comme

$$\varepsilon \text{Log}(a + b) \leq \varepsilon \text{Log } 2 + \varepsilon \text{Log}(a \vee b)$$

et que ε tend vers 0, seul le plus grand terme, ici le premier, compte à la limite et donc

$$\overline{\lim}_{\varepsilon \rightarrow 0} \varepsilon \text{Log } E\zeta_r^\varepsilon(\mathbf{I}) \leq \sup_{u \in \mathbf{B}_\tau \cap \mathcal{S}} \alpha(u).$$

Comme sur \mathcal{S} , α est continue (t. 1) on obtient (3. 1) en faisant tendre τ vers 0.

En fait tout se passe comme si n'avaient été conservés que les n réalisant presque l'égalité dans (3. 7), c'est-à-dire, compte tenu du théorème 1,

les $n_\varepsilon = \frac{1}{\beta(t, x)\varepsilon} + o\left(\frac{1}{\varepsilon}\right)$ où x réalise le sup de $\alpha(t, \cdot)$ sur I .

Démonstration du théorème 2.

A) *Démonstration de a)*

Si $\{t\} \times I$ est contenu dans $(\mathcal{P})^c$, on peut trouver $I', \varepsilon_1 > 0$ et τ (tel que $0 < \tau < t$) avec

i) $B'_\tau = [t - \tau, t] \times I'$ contenu dans $(\mathcal{P})^c$

ii) pour $\varepsilon < \varepsilon_1$, $\frac{1}{\varepsilon}(\{t\} \times I)$ contenu dans $k_\varepsilon B'_{\frac{\tau}{2}}$ avec $k_\varepsilon = 1 + \left\lceil \frac{1}{\varepsilon} \right\rceil$.

La première condition assure que pour tout k et tout n

$$\frac{n}{k} H\left(\frac{k}{n} B'_\tau\right) \geq C_2 > 0$$

la seconde que $\zeta_t^\varepsilon(I)$ est majoré par ρ_{k_ε} où

$$(3.9) \quad \rho_k = \sum_1^N Z_n(kB'_\tau) + \sum_0^N \sum_{y \in \mathbb{R}^n} \chi\left(\lambda_y > k \frac{\tau}{2}\right) + \sum_{N+1}^{+\infty} Z_n(kB'_\tau).$$

On utilise deux majorations différentes pour $n \leq N$ et $n > N$ en suivant (3.2) comme l'indique la figure 4. Reprenant les évaluations (3.3),

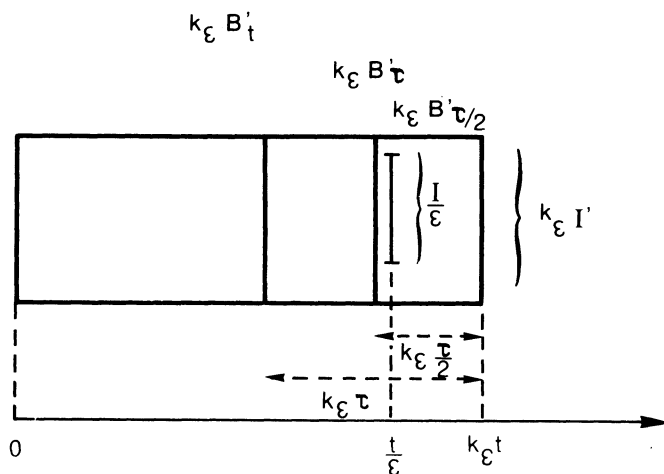


FIG. 4.

(3.4) et (3.5) on trouve, pour un choix de N de type : constante $\times k$, $E\rho_k \leq C_1 k \exp(-C_2 k)$. La série étant convergente, ρ_k tend p. s. vers 0 et, étant à valeurs entières, est p. s. nul à partir d'un certain rang. Il en est donc de même pour $\zeta_t^\varepsilon(I)$ pour ε assez petit.

B) *Démonstration de b).*

Le lemme 1 et l'inégalité de Tcheby'tchev montrent que pour tout $\delta > 0$ $P(\varepsilon \text{ Log } \zeta_t^\varepsilon(I) > \sup_{x \in I} \alpha(t, x) + \delta)$ tend vers 0 quand ε tend vers 0.

Nous voulons montrer maintenant que pour tout $\delta > 0$ on a

$$(3.10) \quad \lim_{\varepsilon \rightarrow 0} P(\varepsilon \text{ Log } \zeta_t^\varepsilon(I) < \sup_{x \in I} \alpha(t, x) - \delta) = 0.$$

Soit D_t la section de \mathcal{S} en t . Comme $I \cap D_t$ est non vide, le sup de $\alpha(t, \cdot)$ sur I est atteint en $x \in I \cap D_t$ et strictement positif. Il suffit de montrer (3.10) pour I inclus dans D_t . $\alpha(t, \cdot)$ est continue sur D_t donc ou bien elle est constante sur I et on peut choisir x dans I , ou bien pour δ' assez petit, inférieur à δ , il existe x_0 dans I tel que $\alpha(t, x_0) = \sup_{x \in I} \alpha(t, x) - \delta'$. Le problème consiste désormais, étant donnés $\delta, I, x_0 \in I$ et $\alpha_0 = \alpha(t, x_0)$ à majorer

$$P_\varepsilon = P(\varepsilon \text{ Log } \zeta_t^\varepsilon(I) < \alpha_0 - \delta).$$

Choisissons I_1 et I_2 , centrés en $\frac{x_0}{\varepsilon}$, de longueurs $2V$ et $2(V + W)$ avec $W > 0$ et inclus dans $\frac{1}{\varepsilon} I$ (ce qui est possible pour $\varepsilon(V + W)$ assez petit) et appliquons (2.1.7) pour n quelconque, t remplacé par $\frac{t}{\varepsilon}$ et W par U (cf. fig. 5).

$$(3.11) \quad \zeta_t^\varepsilon(I) \geq \zeta_{\frac{1}{\varepsilon}}(I_2) \geq Z_n(A) - \sum_{y \in \eta_n(A)} \zeta_{\frac{1}{\varepsilon}}(I_2^c; y), \quad A = \left[\frac{t}{\varepsilon} - U, \frac{t}{\varepsilon} \right] \times I_1$$

ce qui entraîne

$$(3.12) \quad P_\varepsilon \leq P\left(Z_n(A) < \frac{1}{1-\delta} e^{\frac{\alpha_0 - \delta}{\varepsilon}}\right) + P\left(Z_n(A) \geq \frac{1}{1-\delta} e^{\frac{\alpha_0 - \delta}{\varepsilon}}, \quad Z_n(A) - \sum_{y \in \eta_n(A)} \zeta_{\frac{1}{\varepsilon}}(I_2^c; y) < e^{\frac{\alpha_0 - \delta}{\varepsilon}}\right) \leq P\left(Z_n(A) < \frac{1}{1-\delta} e^{\frac{\alpha_0 - \delta}{\varepsilon}}\right) + P\left(Z_n(A) > 0, \sum_{y \in \eta_n(A)} \zeta_{\frac{1}{\varepsilon}}(I_2^c; y) > \delta Z_n(A)\right).$$

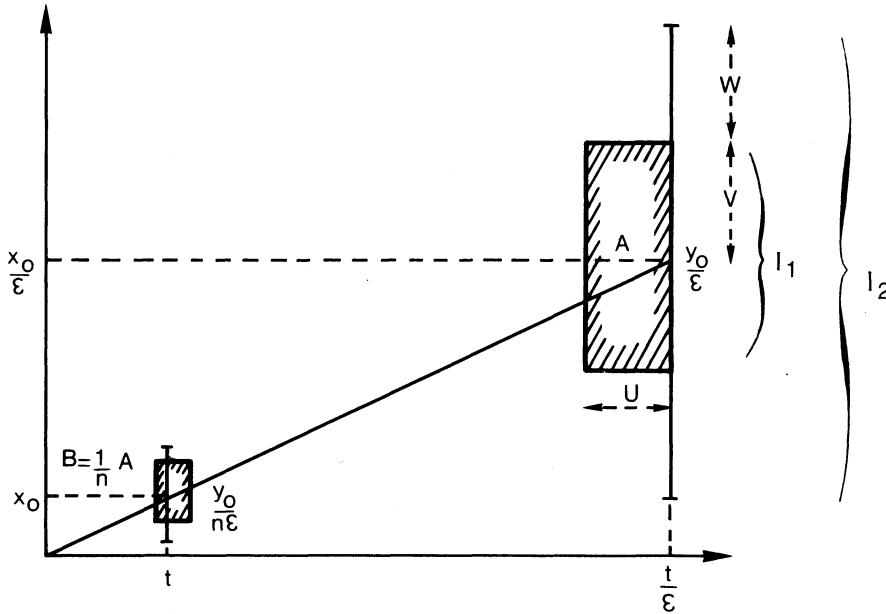


FIG. 5.

Or on peut majorer $P^{\mathcal{B}_n} [Z_n(A) > 0, \sum_{y \in \eta_n(A)} \zeta_{\frac{t}{\epsilon}}(I_2^c; y) > \delta Z_n(A)]$ par

$$\chi(Z_n(A) > 0) \frac{1}{\delta Z_n(A)} \left[\sum_{y \in \eta_n(A)} E^{\mathcal{B}_n} \zeta_{\frac{t}{\epsilon}}(I_2^c; y) \right]$$

et dans cette dernière somme, chacun des termes par $\sup_{x \in I_1, 0 \leq w \leq U} E \zeta_w(I_2^c - x)$.

(On a conditionné par $X_y = \frac{x}{\epsilon}$, $T_y = \frac{t}{\epsilon} - u$ et utilisé le fait que $\zeta_{\frac{t}{\epsilon}}(I_2^c; y)$ a alors conditionnellement la même loi que $\zeta_u(I_2^c - x)$). Mais $E \zeta_u(I_2^c - x)$ est majoré par $\sum_{n \geq 1} Z_n([0, u] \times (I_2^c - x))$ (car pour ce rectangle Z_0 est nul,

0 ne lui appartenant pas) soit pour l'espérance, via la majoration (2.3.8), par $2 \sum_{n \geq 1} \exp(-n \underline{H}(a)) \leq 2 \underline{H}(a)^{-1}$ où on a posé $a = \frac{W}{U}$, supposé assez grand pour que $\underline{H}(a) > 0$.

On a obtenu finalement :

$$(3.13) \quad P\left(\sum_{y \in \eta_n(A)} \zeta_t(I_2^y; y) > \delta Z_n(A)\right) \leq \frac{2}{\delta \underline{H}(a)}$$

qui tend vers 0 dès que $a = \frac{W}{U}$ tend vers l'infini (cf. (2.2.4)).

Il reste à étudier $P\left(Z_n(A) \leq \frac{1}{1-\delta} e^{\frac{\alpha_0 - \delta}{\varepsilon}}\right)$. Via une homothétie, on peut supposer que $y_0 = (t, x_0)$ vérifie $\beta(y_0) = 1$ (cf. théorème 1) et donc $\alpha(y_0) = -h(y_0) > 0$. Choisissons désormais $n = \left\lfloor \frac{1}{\varepsilon} \right\rfloor$. Le rectangle $B = \frac{1}{n}A$ a comme dimensions $\frac{U}{n}, \frac{2V}{n}$ (cf. (3.11)) et contient y_0 dès que $U > t$ et $V > x_0$ (cf. fig. 5). En appliquant (2.3.9), il vient

$$(3.14) \quad P\left(Z_n(A) \leq \frac{1}{1-\delta} e^{\frac{\alpha_0 - \delta}{\varepsilon}}\right) \leq P(\bar{Z}_n^\theta(n(B - y_0)) \leq \delta_\varepsilon)$$

où on a posé $\delta_\varepsilon = \frac{1}{1-\delta} \exp\left(\alpha_0\left(\frac{1}{\varepsilon} - n\right) - \frac{\delta}{\varepsilon} + n\rho|\theta|\right)$ avec $\rho = \max\left(\frac{U}{n}, \frac{V}{n}\right)$. Si on impose à εU et εV de tendre vers 0 quand ε tend vers 0, on a

$$(3.15) \quad \text{Log } \delta_\varepsilon = -\frac{\delta}{\varepsilon} + o\left(\frac{1}{\varepsilon}\right).$$

Nous allons appliquer maintenant la méthode de la génération intermédiaire [11] [6]. L'essentiel des calculs se trouve déjà dans [6], à une seule dimension. Le contexte est ici un peu différent : le rôle joué par la variable t ne permet pas d'espérer une convergence p. s. avec cette méthode; nous utilisons donc une approximation uniforme par la loi normale (Berry-Esséen) et non un théorème limite local, ce qui simplifie les calculs.

Soit s un entier inférieur à n ; qui sera précisé plus tard. On a la décomposition

$$(3.16) \quad \bar{Z}_n^\theta(n(B - y_0)) = \sum_{z \in \mathbb{R}_s} a_z \bar{Z}_{n-s}^\theta(n(B - y_0) - z; z)$$

où

- 1) on a identifié, pour alléger les notations, z et (T_z, X_z)
- 2) on a posé $a_z = \exp((\theta, z) - sL(\theta))$

3) les $\bar{Z}_{n-s}^\theta(\cdot; z)$, pour $z \in \mathbf{R}_s$ sont, conditionnellement à \mathcal{B}_s des copies indépendantes de $\bar{Z}_{n-s}^\theta(\cdot)$.

Il suffit en effet d'écrire
$$\sum_{y \in \mathbf{R}_n} = \sum_{z \in \mathbf{R}_s} \sum_{y \in \mathbf{R}_{n-s}(z)}$$
 et d'appliquer (2.1.1).

Conditionnellement à \mathcal{B}_s , la v. a. $\bar{Z}_n^\theta(n(\mathbf{B} - y_0))$ est donc une combinaison linéaire de v. a. i. Elles n'ont pas même loi mais nous pouvons contrôler leurs queues. Nous allons les centrer conditionnellement. Définissons pour $z \in \mathbf{R}_s$

$$(3.17) \quad b_z = \bar{\mu}_{n-s}^\theta(n(\mathbf{B} - y_0) - z) \\ V_z = \bar{Z}_{n-s}^\theta(n(\mathbf{B} - y_0) - z; z) - b_z \quad (E^{\mathcal{B}_s} V_z = 0 \text{ d'après 2.3.5}).$$

Dans ce qui suit, les sommations non précisées sont faites sur $z \in \mathbf{R}_s$; Si p est un réel positif, destiné à tendre vers 0, on a :

$$(3.18) \quad P(\bar{Z}_n^\theta(n(\mathbf{B} - y_0)) \leq \delta_\varepsilon) \leq P_1 + P_2 \\ P_1 = P\left(\sum a_z b_z \leq p\right) \\ P_2 = P\left(\sum a_z V_z < \delta_\varepsilon - p\right).$$

Si s tend vers l'infini et $\frac{s}{n}$ tend vers 0, une application de Berry-Esséen permet de majorer P_1 . Cette majoration n'est utile, autrement dit $\bar{Z}_n^\theta(n(\mathbf{B} - y_0))$ n'est proche de son espérance conditionnelle que s'il y a suffisamment d'individus dans $n\mathbf{B}$: la majoration de P_2 se fait sous l'hypothèse $\{t\} \times I \subset \mathcal{S}$ qui entraîne $h(y_0) < 0$. L'hypothèse (H5) sert dans la majoration de P_2 mais pas dans celle de P_1 .

1) ÉTUDE DE P_1 .

Minorons b_z grâce au lemme suivant

LEMME 2. — Soit ν une probabilité sur \mathbb{R}^2 , centrée, de matrice de covariance Γ de rang 2, de troisièmes moments finis. $C_n = L_n \times M_n$ est une suite de rectangles contenant 0 et tels que, quand n tend vers l'infini, $\frac{|L_n|}{\sqrt{n}}$ et $\frac{|M_n|}{\sqrt{n}}$ restent minorés par des nombres strictement positifs. k_n est

une suite d'entiers telle que $\frac{k_n}{n}$ tende vers 1. Alors il existe deux constantes K_1 et $K_2 > 0$ telles que pour tout z , et tout n assez grand

$$(3.19) \quad v^{*k_n}(C_n - z) \geq K_1 - K_2 \frac{\|z\|}{\sqrt{n}}$$

où $\| \cdot \|$ désigne la norme euclidienne dans \mathbb{R}^2 .

Démonstration du lemme 2. — Soit Ψ la loi $\mathcal{N}(0, \Gamma)$. D'après Berry-Esséen, on peut écrire

$$v^{*k_n}(C_n - z) = O\left(\frac{1}{\sqrt{k_n}}\right) + \Psi\left(\frac{1}{\sqrt{k_n}}(C_n - z)\right).$$

La relation $|\Psi(A + u) - \Psi(A)| = O_{\Gamma}(\|u\|)$ permet d'écrire

$$v^{*k_n}(C_n - z) \geq O\left(\frac{1}{\sqrt{n}}\right) + \Psi\left(\frac{1}{\sqrt{k_n}}C_n\right) - K_2 \frac{\|z\|}{\sqrt{n}}.$$

Par hypothèse, $\frac{1}{\sqrt{k_n}}C_n$ contient un rectangle fixe contenant 0, donc $\Psi\left(\frac{1}{\sqrt{k_n}}C_n\right)$ reste minoré, ce qui achève la démonstration.

Appliquons ce lemme à $k_n = n - s$ (désormais $\frac{s}{n} = o(1)$), $v = \bar{\mu}^{\theta}$ et $C_n = n(B - y_0)$. C_n contient 0 car B contient y_0 . L'hypothèse sur les dimensions sera vérifiée si $\frac{U}{\sqrt{n}}$ et $\frac{2V}{\sqrt{n}}$ restent minorés, donc si $\sqrt{\varepsilon}U$ et $\sqrt{\varepsilon}V$ restent minorés quand ε tend vers 0.

On peut résumer les contraintes sur U, V, W par :

$$(3.20) \quad \begin{array}{lll} \sqrt{\varepsilon}U & \text{et} & \sqrt{\varepsilon}V > \text{cste} > 0 & \varepsilon(V + W) \rightarrow 0 & (\text{cf. (3.11)}) \\ \frac{W}{U} \rightarrow \infty & (\text{cf. (3.14)}) & & \varepsilon U \rightarrow 0 & (\text{cf. (3.16)}) \end{array}$$

ce qu'il est toujours possible de réaliser.

Dans ces conditions on a $b_z = \bar{\mu}_{n-s}^{\theta}(n(B - y_0) - z) \geq K_1 - K_2 \frac{\|z\|}{\sqrt{n}}$ et en remarquant que $\Sigma a_z = \bar{Z}_s^{\theta}(\mathbb{R}^+ \times \mathbb{R})$ on obtient :

$$P_1 = P(\Sigma a_z b_z < p) \leq P(K_1 \bar{Z}_s^{\theta}(\mathbb{R}^+ \times \mathbb{R}) \leq 2p) + P(K_2 \Sigma \|z\| a_z > p\sqrt{n})$$

θ est fixe (dual de y_0), donc $\bar{Z}_s^{\theta}(\mathbb{R}^+ \times \mathbb{R})$ en temps que martingale positive ([J], p. 248) a une limite p. s. quand s tend vers l'infini. Il est facile de

montrer que sous (H2) (qui entraîne la non-extinction p. s.) cette limite n'a pas de masse en 0 donc le premier terme de la majoration de P_1 tend vers 0 dès que p tend vers 0. Pour finir, on applique Tchebychev :

$$P(\Sigma \|z\| a_z > Ap\varepsilon^{-1/2}) \leq \frac{\varepsilon^{1/2}}{Ap} E(\Sigma \|z\| a_z)$$

et comme $E(\Sigma \|z\| a_z) = \int \|z\| \mu_s^\theta(dz) = 0(s)$ on conclut

$$(3.21) \quad P_1 = o(1) + 0(p^{-1}\varepsilon^{1/2}s) \quad (s = o(\varepsilon^{-1}), p = o(1)).$$

2) ÉTUDE DE P_2 .

$$(3.22) \quad P_2 = P(\Sigma a_z V_z \leq \delta_\varepsilon - p) \leq P\left(|\Sigma a_z V_z| > \frac{p}{2}\right), \quad \text{dès que } p > 2\delta_\varepsilon$$

si Φ est la fonction croissante définie sur \mathbb{R}^+ par

$$(3.23) \quad \Phi(u) = u^2\chi(u < 1) + (2u - 1)\chi(u \geq 1), \quad \text{si } p < 2$$

(3.22) entraîne

$$(3.23) \quad P_2 \leq \frac{4}{p^2} E\Phi(|\Sigma a_z V_z|), \quad \text{si } p < 2$$

Φ est convexe et vérifie $\Phi(2x) \leq 4\Phi(x)$ pour tout x de \mathbb{R}^+ ; les a_z sont \mathcal{B}_s mesurables et les V_z conditionnellement à \mathcal{B}_s sont indépendantes et centrées. D'après le théorème 1.1 de [9] il existe donc une constante C_Φ ne dépendant que de Φ telle que

$$E^{\mathcal{B}_s}\Phi(|\Sigma a_z V_z|) \leq C_\Phi E^{\mathcal{B}_s}\Phi(\sqrt{\Sigma(a_z V_z)^2}).$$

La fonction Φ_1 définie par $\Phi_1(x) = \Phi(\sqrt{x})$ est concave donc sous-additive et par conséquent :

$$(3.24) \quad E^{\mathcal{B}_s}\Phi(|\Sigma a_z V_z|) \leq C_\Phi \Sigma E^{\mathcal{B}_s}\Phi(a_z | V_z|).$$

Pour étudier $E^{\mathcal{B}_s}\Phi(a_z | V_z|)$ on remarque d'après (3.16) et (3.17) que $|V_s| \leq Y_z$ où conditionnellement à \mathcal{B}_s les Y_z sont des v. a. i. et de même loi que $Y = 1 + \sup_n \bar{Z}_n^\theta(\mathbb{R}^+ \times \mathbb{R})$. (H5) assure [6] que Y est p. s. finie de loi Q vérifiant

$$(3.25) \quad \int g_{k-1}(u)Q(du) < \infty.$$

La majoration de V_z entraîne, pour toute f croissante de \mathbb{R}^+ dans \mathbb{R} ,

$$(3.26) \quad E^{\mathcal{A}_s} f(|V_z|) \leq \int f(u) Q(du)$$

ce qui joint à (3.24) et (3.23) permet la majoration

$$(3.27) \quad P_2 \leq \frac{4C_\Phi}{p^2} E\Sigma \int \Phi(a_z u) Q(du) = \frac{4C_\Phi}{p^2} \iint \Phi(au) Q(du) v_s(da)$$

où $v_s(da)$ est la mesure image de $\mu^{*s}(dz)$ par l'application

$$\varphi : z \mapsto a_z = \exp((\theta, z) - sL(\theta)).$$

Comme $av_s(da)$ est l'image par φ de μ_s^θ , on a $\int av_s(da) = 1$. Pour évaluer l'intégrale double de (3.27), on commence par majorer $\Phi(au)$ par $\Phi_2(au) = 2[(au)^2 \wedge (au)]$ (cf. (3.23)) et décomposer $\mathbb{R}^+ \times \mathbb{R}^+$ en

$$\Delta_1 = \{u > M\}, \quad \Delta_2 = \left\{u < M, a < \frac{1}{M}\right\}, \quad \Delta_3 = \left\{u < M, a > \frac{1}{M}\right\},$$

où $M > 1$ sera fixé plus tard. On a successivement :

$$(3.28) \quad \begin{aligned} \text{sur } \Delta_1 \quad \Phi_2(au) &\leq 2au & \text{et} & \quad \iint_{\Delta_1} \leq 2 \int_{u>M} u Q(du) \\ \text{sur } \Delta_2 \quad \Phi_2(au) &\leq 2a^2u^2 \leq 2 \frac{au^2}{M} & \text{et} & \quad \iint_{\Delta_2} \leq 2 \int_{u<M} \frac{u^2}{M} Q(du) \\ \text{sur } \Delta_3 \quad \Phi_1(au) &\leq 2au & \text{et} & \quad \iint_{\Delta_3} \leq 2 \left[\int_{a>\frac{1}{M}} av_s(da) \right] \left[\int u Q(du) \right] \end{aligned}$$

$$\iint_{\Delta_1 \cup \Delta_2} \leq 2 \int \left(u \wedge \frac{u^2}{M} \right) Q(du) = 0 ((\text{Log } M)^{-k+1})$$

$$\text{d'après (3.2.5) et } \sup_u \frac{u \wedge \frac{u^2}{M}}{g_{k-1}(u)} = \sup_u \frac{1 \wedge \frac{u}{M}}{[1 + (\text{Log}^+ u)^{k-1}]} = 0 ((\text{Log } M)^{-k+1})$$

$$(3.29) \quad \iint_{\Delta_3} = 0 \left[\int_{a>\frac{1}{M}} av_s(da) \right]$$

soit $l > 0$ et $M = e^{ls}$. On a $\varphi^{-1}\left(a > \frac{1}{M}\right) = sD_l$ avec $D_l = \{z: (\theta, z) - L(\theta) > -l\}$ et par conséquent

$$\int_{a > \frac{1}{M}} av_s(da) = \mu_s^\theta(sD_l)$$

μ^θ a comme barycentre y_0 , vérifiant $h(y_0) > 0$. Si l est choisi tel que $0 < l < -h(y_0)$, D_l ne contient pas y_0 et donc d'après (2.3.8) il existe $\gamma > 0$ tel que $\mu_s^\theta(sD_l) \leq e^{-\gamma s}$, ce qui donne

$$(3.30) \quad \iint_{\Delta_1 \cup \Delta_2} = O(s^{-k+1}) \quad \iint_{\Delta_3} = O(e^{-\gamma s})$$

et, joint à (3.27)

$$(3.31) \quad P_2 = O(p^{-2}s^{-k+1}).$$

Les contraintes en p, s pour que P_1 et P_2 tendent vers 0 avec ε sont :

$$p^{-1} \exp\left(-\frac{\delta}{\varepsilon} + o\left(\frac{1}{\varepsilon}\right)\right) < \frac{1}{2} \quad p \rightarrow 0, \quad s\varepsilon \rightarrow 0, \quad s \rightarrow \infty$$

$$sp^{-1}\varepsilon^{1/2} \rightarrow 0 \quad s^{-k+1}p^{-2} \rightarrow 0.$$

On peut prendre par exemple $p = \varepsilon^{\frac{k-1}{4(k+1)}}$ $s = \left[\varepsilon^{-\frac{1}{k+1}}\right]$.

IV. DISCUSSION

1. Deux cas particuliers.

Nous supposons ici que l'instant de naissance de tous les enfants coïncide avec celui de la mort du père (processus dit de Sevastyanov [10], p. 8).

a) Absence de marche aléatoire dans l'espace :

c'est le processus de Bellman-Harris classique ; on a

$$\mu(dt, dx) = mG(dt)\delta_0(x) \quad (5), \quad L(\theta_1, \theta_2) = \text{Log } m + L_G(\theta_1)$$

$\alpha(t, x)$ se simplifie en αt avec α donné par la formule (1.2). La relation (2.2.7) devient

$$(4.1.1) \quad \text{Log} \int_0^\infty e^{-\alpha t} mG(dt) = 0.$$

(5) δ_a désigne la masse de Dirac au point a .

On retrouve le résultat classique (1.1) sur le paramètre de Malthus [10].

b) Absence de marche aléatoire dans le temps :

c'est le processus de branchement avec marche aléatoire à temps discret [4] [11] ; on a, si F est la mesure d'intensité spatiale :

$$(4.1.2) \quad \alpha(t, x) = - t h_F \left(\frac{x}{t} \right)$$

(qui est à rapprocher de [4], p. 676). Enfin la formule (2.2.7) devient :

$$(4.1.3) \quad \frac{\partial \alpha}{\partial t} = L_F \left(- \frac{\partial \alpha}{\partial x} \right).$$

Ces cas particuliers permettent en retour d'éclairer le cas général. Pour θ_2 fixé, la mesure $\mu_{\theta_2}(dt) = \int_{\mathbb{R}} e^{\theta_2 x} \mu(dt, dx)$ ne charge pas $\{0\}$ et a un support inclus dans \mathbb{R}^+ donc il existe un unique θ_1 tel que $\text{Log} \int_0^\infty e^{\theta_1 t} \mu_{\theta_2}(dt) = 0$. c'est-à-dire $L(\theta_1, \theta_2) = 0$ ($-\theta_1$ est aussi la pente de la tangente menée de $(0, 0)$ au graphe de $h_{\mu_{\theta_2}}$). Ceci permet de définir une fonction \tilde{L} définie partout telle que :

$$L(\theta_1, \theta_2) = 0 \Leftrightarrow \theta_1 = - \tilde{L}(\theta_2).$$

En vertu du théorème des fonctions implicites $\left(\frac{\partial L}{\partial \theta} > 0 \right)$ d'après (H3) et [2] prop. (9.7), \tilde{L} est C^∞ et on vérifie $\tilde{L}'' > 0$ donc \tilde{L} est convexe. La formule (2.2.7) devient

$$(4.1.4) \quad \frac{\partial \alpha}{\partial t} = \tilde{L} \left(- \frac{\partial \alpha}{\partial x} \right).$$

Elle est analogue à (4.1.3) et entraîne, si \tilde{h} est la duale convexe de \tilde{L}

$$(4.1.5) \quad \alpha(t, x) = - t \tilde{h} \left(\frac{x}{t} \right).$$

Par homogénéité, le sup de $\alpha(t, \cdot)$ sur \mathbb{R} est atteint en $t \tilde{x}$ où \tilde{x} réalise l'inf de \tilde{h} , et ce sup vaut $- t \tilde{h}(\tilde{x}) = t \tilde{L}(0)$ par dualité. D'après le théorème 2, la population totale est donc de l'ordre de $\exp \left(\frac{1}{\varepsilon} (t \tilde{L}(0) + o(1)) \right)$. Or par construction $\tilde{L}(0)$ est le paramètre de Malthus d'un processus de Bellman-Harris où $\mu_0(dt)$ jouerait le rôle de $mG(dt)$. Dans l'échelle utilisée ($\varepsilon \text{ Log}$),

à l'instant t la population totale est donc du même ordre que celle située « près de » $t\tilde{x}$; c'est aussi celle d'un processus non spatial régi par la marginale de μ en t .

2. Affaiblissement des hypothèses.

- a) Nous avons déjà dit que (H1) est nécessaire pour que $\mathcal{S} \neq \emptyset$.
 b) Nous conjecturons que (H2) utilisée en (2.1.7) peut être affaiblie en

$$(4.2.1) \quad \mathbf{P}(\xi([0, \lambda] \times \mathbb{R}) \geq 1) > 0$$

le théorème 2 restant valable conditionnellement à la non-extinction. En effet soit $\tau = \inf \{ t : \xi([0, t] \times \mathbb{R}) > 1 \}$ (âge du père à la naissance de son 1^{er} enfant). Seul le cas $0 < \mathbf{P}(\lambda \leq \tau) < 1$ reste à traiter. Soit $\lambda^+ = \max(\lambda, \tau)$ et $\zeta_t^+(\mathbf{I})$ la v. a. obtenue en remplaçant dans (2.1.3) λ par λ^+ . Il est clair que $\zeta_t^+(\mathbf{I}) \geq \zeta_t(\mathbf{I})$ d'où

$$\overline{\lim} \varepsilon \text{Log } \zeta_t^\varepsilon(\mathbf{I}) \leq \lim \varepsilon \text{Log } \zeta_t^{+\varepsilon}(\mathbf{I}) = \sup_{x \in \mathbf{I}} \alpha(t, x).$$

Soit $\zeta_t^-(\mathbf{I})$ le nombre d'individus de la nouvelle population, en vie à l'instant t , situés dans \mathbf{I} et vérifiant $\lambda_y^+ = \lambda_y$. On a évidemment $\zeta_t(\mathbf{I}) \geq \zeta_t^-(\mathbf{I})$ (on a supprimé dans la population de référence née avant l'instant t les individus n'ayant que des enfants posthumes). Soit $\eta_n^-(A)$ le sous-ensemble de $\eta_n(A)$ formé des individus dont tous les descendants en vie à t vérifient $\lambda_y^+ = \lambda_y$. D'après le principe d'exclusion (2.1.6), tout se passe comme si dans $\eta_n(A)$ on avait conservé ou non chaque individu, indépendamment les uns des autres (conditionnellement à \mathcal{B}_n). Ceci suggère une loi des grands nombres (conditionnelle). Si toutes les probabilités de conservation étaient les mêmes (p_n) on aurait $\frac{Z_n^-(A)}{Z_n(A)} \simeq p_n$ et la démonstration à partir de (3.11) conduirait à

$$\underline{\lim} \zeta_t^\varepsilon(\mathbf{I}) \geq \sup_{x \in \mathbf{I}} \alpha(t, x)$$

pourvu que $\varepsilon \text{Log } p_n$ tende vers 0. La difficulté réside dans la dépendance de ces probabilités par rapport à l'instant de naissance : si cet instant est noté $\frac{t}{\varepsilon} - u$, cette probabilité (conditionnellement à \mathcal{B}_n) vaut $E p^{\varepsilon u(\mathbb{R})}$ où $p = \mathbf{P}(\lambda \leq \tau)$. Mais conditionnellement à la non-extinction, elle est équivalente à $Q(q)e^{-\alpha u}$ quand u tend vers l'infini ([1], p. 162, $Q(q) > 0$, α paramètre de Malthus global). Notre conjecture s'appuie alors sur le fait que

$u \leq U$, εU tend vers 0 et $\sqrt{\varepsilon}U$ reste minoré. Il ne semble pas possible de descendre en deçà de (4.2.1). On peut en effet construire un processus de branchement tel que p. s. $\lambda = \frac{\tau}{2} = l$ (l fixé déterministe). Alors pour tout t il existe toujours une suite ε_k tendant vers 0 telle que $\zeta_t^{\varepsilon_k}(\mathbb{R}) = 0$ pour tout k .

c) (H3) est fondamentale. Seul 1) pourrait sans doute être affaibli en : il existe un voisinage de 0 dans \mathbb{R}^2 sur lequel $\exp L(\cdot) < \infty$.

d) (H4) est utilisée dans la démonstration du lemme 1. A première vue elle peut sembler artificielle, le résultat final ne faisant intervenir que la fonction α , qui ne dépend pas de la forme de G (remarque également valable pour (H2)). Il suffit pour se convaincre de son utilité, d'étudier le cas extrême de la durée de vie infinie. Dans un tel cas, on a évidemment

$$\overline{\lim} \varepsilon \text{Log } \zeta_t^\varepsilon(\mathbb{I}) \leq \sup_{[0,t] \times \mathbb{I}} \alpha(y)$$

et on minore $\zeta_t^\varepsilon(\mathbb{I})$ par $\eta_n(A')$ où A' est un certain rectangle construit à partir de y' réalisant le sup, avec naturellement $n = \left\lfloor \frac{1}{\beta(y')\varepsilon} \right\rfloor$, d'où la formule :

$$(4.2.2) \quad \lim \varepsilon \text{Log } \zeta_t^\varepsilon(\mathbb{I}) = \sup_{[0,t] \times \mathbb{I}} \alpha(y).$$

Dans un cas intermédiaire, si $\eta_0 = \sup \{ \eta : \hat{G}(\eta) < \infty \}$ vérifie $0 < \eta_0 < \infty$, nous conjecturons l'existence d'une limite

$$(4.2.3) \quad \lim \varepsilon \text{Log } \zeta_t^\varepsilon(\mathbb{I}) = \sup_{x \in \mathbb{I}} \alpha_{\eta_0}(t, x)$$

$$\alpha_{\eta_0}(t, x) = \sup_{0 < t' < t} [\alpha(t', x) - \eta_0(t - t')].$$

Cette conjecture s'appuie sur l'idée suivante. Si la durée de vie est exponentielle de paramètre η_0 , construisons un nouveau processus de branchement à 2 types d'individus. Ceux du 1^{er} type donnent à la fois des individus du 1^{er} type de la manière habituelle (ξ) et au bout d'un temps T fixé (suffisant pour vérifier (4.2.1)), s'ils ne sont pas déjà morts, meurent et avec une probabilité $e^{-\eta_0}$ donnent un enfant de type 2 au même endroit. Ceux de type 2 vivent un temps unité au bout duquel ils donnent avec probabilité $e^{-\eta_0}$ un autre individu au même endroit. Si on s'intéresse à l'effectif total (types 1 et 2) dans \mathbb{I} à l'instant t les méthodes de processus multitypes [14] conduisent à (4.2.3). D'autre part nous conjecturons que cet effectif total a même loi que celui du processus unitype initial.

e) (H5) est du genre « $x \text{Log } x$ », hypothèse usuelle dans les résultats sur les processus de branchement.

3. Conjecture sur les âges.

Soit $(t, x) \in \mathcal{S}$. On a vu que pour la population en vie à l'instant t autour de x , l'essentiel venait des générations de numéro $\varepsilon^{-1}\beta(t, x)^{-1} + o(\varepsilon^{-1})$. Tout semble donc se passer comme si le saut moyen dans l'espace-temps, correspondant à une génération était

$$\frac{\varepsilon^{-1}t}{\varepsilon^{-1}\beta(t, x)^{-1} + o(\varepsilon^{-1})} = t\beta(t, x) + o(1),$$

$$\frac{\varepsilon^{-1}x}{\varepsilon^{-1}\beta(t, x)^{-1} + o(\varepsilon^{-1})} = x\beta(t, x) + o(1).$$

Or d'après la dualité, si $y = (t, x)$ et $(\theta_1, \theta_2) = \text{grad } h(\beta(y)y)$ on a

$$t\beta(t, x) = \int \tau \exp [\theta_1\tau + \theta_2\xi - L(\theta_1, \theta_2)] \mu(d\tau, d\xi)$$

$$x\beta(t, x) = \int \xi \exp [\theta_1\tau + \theta_2\xi - L(\theta_1, \theta_2)] \mu(d\tau, d\xi).$$

Ceci suggère que pour la plupart des individus en vie à l'instant t , autour de x , la loi empirique du saut, calculée le long de leur ligne généalogique est voisine de

$$(\tau, \xi) \mapsto \exp [\theta_1\tau + \theta_2\xi - L(\theta_1, \theta_2)] \mu(d\tau, d\xi).$$

De tels résultats ont été démontrés rigoureusement pour le processus à temps discret unitype [13] et multitype [14]. Le fait que $\alpha(t, x) = \theta_1 t + \theta_2 x$ joue le rôle d'un paramètre de Malthus local permet de conjecturer en plus que la répartition des âges à l'instant t autour de x converge vers la loi sur \mathbb{R}^+

$$\tau \mapsto \frac{e^{\theta_1\tau} G(\cdot, +\infty) d\tau}{\int_{\mathbb{R}^+} e^{\theta_1 u} G(\cdot, +\infty) du}$$

Si on se place en $t\tilde{x}$ (cf. IV, 1.b) on a $\theta_1 = \tilde{L}(0)$ et on retrouve la distribution stable des âges pour un processus non spatial régi par la marginale de μ en t [10].

4. Processus non-homogène.

Revenons sur l'intérêt de l'introduction de ε et donc sur la différence d'optique par rapport à Biggins [7] déjà évoquée dans l'introduction.

Dans le cas présent on peut effectivement se passer de ε car α est 1-homogène. Ceci est dû au fait que le processus de reproduction ne dépend pas de la position du père dans l'espace-temps. On peut se libérer de cette restriction en supposant qu'un père en (σ, a) a des enfants répartis suivant un processus ponctuel d'intensité $\mu(\sigma, a; \cdot)$. L'analogue pour un seul individu serait une chaîne de Markov à valeurs dans $\mathbb{R}^+ \times \mathbb{R}$. La théorie des grandes déviations dans ce cas a été construite dans [3] et dans [15]. A la famille $(\mu(\sigma, a; \cdot))_{\sigma, a}$ on associe $(L(\sigma, a; \cdot))_{\sigma, a}$ et $(h(\sigma, a; \cdot))_{\sigma, a}$ et à tout $T > 0$ et tout chemin $\varphi[0, T] \rightarrow \mathbb{R}^+ \times \mathbb{R}$ de classe \mathcal{C}^1 par morceaux on associe l'intégrale d'action

$$I(T, \varphi) = \int_0^T h(s, \varphi(s); \dot{\varphi}(s)) ds.$$

Dans un travail ultérieur nous montrerons qu'on peut encore définir une ε -log densité α , l'analogue de (2.2.6) étant $-\inf_T \inf_{\varphi \in \Gamma_T(t, x)} I(T, \varphi)$ où $\Gamma_T(t, x)$ est un sous-ensemble convenable de $\{\varphi : \varphi(T) = (t, x)\}$. On peut attendre comme analogue de (2.2.7) une inéquation aux dérivées partielles portant sur $L\left(t, x; -\frac{\partial \alpha}{\partial t}, \frac{\partial \alpha}{\partial x}\right)$ permettant éventuellement de construire des schémas numériques d'approximation de la solution α . Nous envisagerons aussi le cas où le processus de reproduction dépend de l'effectif de la population autour du père. L'équation (2.2.7) est de type Hamilton-Jacobi, de hamiltonien L . Le lagrangien h est associé à un problème variationnel qui trouve ici une illustration stochastique, les chemins φ étant constitués par les trajectoires le long des lignes généalogiques.

REMERCIEMENTS

Nous remercions le Professeur Ruget pour les conversations stimulantes que nous avons pu avoir avec lui.

BIBLIOGRAPHIE

- [1] K. B. ATHREYA, P. NEY, *Branching processes* (Springer-Verlag), Berlin, 1972.
- [2] R. AZENCOTT, Cours de l'École d'Été de Saint-Flour. *Lecture Notes in Math.*, n° 774, Springer, 1978.
- [3] R. AZENCOTT, G. RUGET, Mélanges d'équations différentielles et grands écarts à la loi des grands nombres. *Zeits. für Wahr.* t. **38**, 1977, p. 1-54.
- [4] J. D. BIGGINS, Chernoff's theorem in the branching random walk. *Journ. of Appl. Prob.*, t. **14**, 1977, p. 630-636.
- [5] J. D. BIGGINS, The asymptotic shape of the branching random walk. *Adv. in Appl. Prob.*, t. **10**, 1978, p. 62-84.

- [6] J. D. BIGGINS, Growth rates in the branching random walk. *Zeits. für Wahr.* t. **48**, 1979, p. 17-34.
- [7] J. D. BIGGINS, Spatial spread in branching processes, in « Biological growth and spread ». *Lecture Notes in Biomath.*, n° 38, Springer, 1979.
- [8] J. D. BIGGINS, *Private communication*, 1980.
- [9] D. L. BURKHOLDER, B. J. DAVIS, R. F. GUNDY, *Integral inequalities for convex functions of operators on martingales*, 6th Berkeley symposium, t. **II**, 1970.
- [10] P. JAGERS, *Branching processes with biological applications*. Wiley, London, 1979.
- [11] N. KAPLAN, S. ASMUSSEN, Branching random walks II. *Stochastic Proc. and their Appl.*, t. **4**, 1976, p. 15-31.
- [12] K. MATTHES, J. KERSTAN, J. MECKE, *Infinitely divisible point processes*. Wiley Chichester, 1978.
- [13] A. ROUAULT, Lois empiriques pour les processus de branchement spatiaux homogènes supercritiques. *Note aux C. R. A. S.*, t. **292**, 1981, p. 933.
- [14] A. ROUAULT, *Les processus de branchement multitypes spatiaux dans l'asymptotique des grandes déviations*, 1981, (à paraître).
- [15] A. D. VENTSEL, Rough limit theorems on large deviations for Markov stochastic processes. *Theory of Prob. and its Appl.*, t. **21**, n° 2, 1976, p. 227-242 et t. **21**, n° 3, 1976, p. 499-512.

(Manuscrit reçu le 1^{er} octobre 1982)

E. Lois empiriques dans les processus de branchements spatiaux
homogènes supercritiques.

PROBABILITÉS. — *Lois empiriques dans les processus de branchement spatiaux homogènes supercritiques.* Note (*) de **Alain Rouault**, transmise par Robert Fortet.

Pour chaque réalisation ω d'un processus de branchement spatial homogène supercritique, on tire uniformément un individu dans la n -ième génération. Ceci probabilise les généalogies. On montre, pour presque tout ω de l'ensemble de survie, une majoration de grandes déviations pour la loi empirique du processus de reproduction, ce qui permet une description à vitesse moyenne fixée.

For each realization ω of a supercritical spatially homogeneous branching process, a random sampling among the n -th generation yields a probability law on the set of genealogies. For almost every ω of the survival set, we show a large deviation upper bound for the empirical law of the offspring process, which allows a description when mean velocity is fixed.

I. POSITION DU PROBLÈME. — Soit \mathcal{M} l'ensemble des mesures de Radon positives sur \mathbb{R} , $\mathcal{B}(\mathcal{M})$ la tribu borélienne pour la topologie vague. L'ensemble M (resp. \dot{M}) des mesures ponctuelles (resp. mes. ponct. simples) sur \mathbb{R} est un fermé (resp. borélien) de \mathcal{M} . Soit $\mathcal{B}(M)$ la tribu trace. Soit P une probabilité sur $(M, \mathcal{B}(M))$ portée par \dot{M} et vérifiant :

$$(1) \quad \int_M \Phi(\mathbb{R}) \log \Phi(\mathbb{R}) P(d\Phi) < \infty.$$

Son intensité μ , mesure positive sur \mathbb{R} , est supposée sans atome et vérifiant :

$$(2) \quad 1 < m = \mu(\mathbb{R}) < \infty \quad \text{d'où} \quad P(\Phi : \Phi(\mathbb{R}) < \infty) = 1.$$

On note $\mu_0 = m^{-1} \mu$.

$$(3) \quad L(\theta) = \log \int e^{\theta x} \mu(dx) < \infty \quad \text{pour} \quad |\theta| < \theta_M.$$

Il existe une indexation mesurable $(x, \Phi) \rightarrow i(x, \Phi)$ de l'ensemble $C = \{(x, \Phi) \in \mathbb{R} \times M; \Phi(\{x\}) > 0\}$ fermé de $\mathbb{R} \times M$ dans \mathbb{N} [7].

Soit $(\Omega, \mathcal{A}, P) = \bigotimes_{i,j=1}^{\infty} (M, \mathcal{B}(M), P)$, $\mathcal{F}_n = \sigma\{\omega_{i,j}, i \leq n, j \in \mathbb{N}\}$, $Z_0 = \delta_0$ masse de Dirac en 0 :

$$Z_1^\omega = \omega_{1,0},$$

pour $n \geq 2$:

$$Z_n^\omega = \int \tau_x \omega_{n, i(x, Z_{n-1}^\omega)} Z_{n-1}^\omega(dx),$$

où τ_x est l'opérateur translation par x .

$Z_n^\omega \in \dot{M}$ p. s. car μ est sans atome, Z_n est \mathcal{F}_n -mesurable. Si Z_{n-1}^ω décrit les positions de la $n-1$ -ième génération, Z_n^ω s'obtient en faisant de chaque point de Z_{n-1}^ω l'origine d'une mesure ponctuelle tirée suivant P , indépendamment de \mathcal{F}_{n-1} . (2) assure que le processus est supercritique non-explosif. Soit $S = \{\omega : Z_n^\omega(\mathbb{R}) \rightarrow +\infty\}$, $P(S) > 0$. Si h_0 désigne la transformée de Cramer [3] de μ_0 et si a vérifie $(1) h_0(a) < \log m$ et $e_0 = \int x \mu_0(dx) < a$, alors

p. s. sur S :

$$(4) \quad n^{-1} \log Z_n^\omega(n[a, a + \varepsilon]) \rightarrow -h_0(a) + \log m \quad \text{pour tout } \varepsilon > 0 [2].$$

Il y a donc un très grand nombre d'individus de la n -ième génération de « vitesse moyenne » dans $[a, a + \varepsilon]$. La question est de savoir quelle a été, en moyenne, l'histoire de leurs ancêtres, notamment comment ils se sont reproduits.

Pour $\omega \in S$ et x tel que $Z_n^\omega(\{x\}) = 1$ on définit $\pi^\omega(x) \in \mathbb{R}$ par :

$$Z_{n-1}^\omega(\{\pi^\omega(x)\}) = 1 \quad \text{et} \quad \tau_{\pi^\omega(x)} \omega_{n, i(\pi^\omega(x), Z_{n-1}^\omega)}(\{x\}) = 1$$

et on note $r^\omega(x) = \omega_{n, i(\pi^\omega(x), Z_{n-1}^\omega)}$.

$\pi^\omega(x)$ donne la position du père de x et $r^\omega(x)$ la position de x et de ses frères par rapport à leur père.

Soit $\pi_0^\omega =$ identité, $\pi_k^\omega = \pi_{k-1}^\omega \circ \pi^\omega$, $k \geq 1$ et :

$$Y_k^\omega(x) = (\pi_{n-k}^\omega(x) - \pi_{n-k+1}^\omega(x), r^\omega \circ \pi_{n-k}^\omega(x)) \in C,$$

$$\Gamma_n^\omega = \{(\gamma_1, \dots, \gamma_n) \in C^n : \exists x : Z_n^\omega(\{x\}) = 1 \quad \text{et} \quad \forall i \leq n, \gamma_i = Y_i^\omega(x)\},$$

désigne l'ensemble des n -généalogies. Son cardinal vaut $Z_n^\omega(\mathbb{R})$. Soit \mathcal{P}_n^ω la loi uniforme sur Γ_n^ω .

II. RÉSULTATS. — Sur C on définit $\mathcal{C}_p(dx, d\Phi) = \Phi(dx) P(d\Phi)$ mesure de Campbell associée à P ([5], [7]).

LEMME. — Pour toute φ continue bornée de C^n dans \mathbb{R} :

$$(5) \quad E \int_{\mathbb{R}} \varphi(Y_1^\omega(x), \dots, Y_n^\omega(x)) Z_n^\omega(dx) = \int_{C^n} \varphi(\gamma_1, \dots, \gamma_n) \mathcal{C}_p(d\gamma_1) \dots \mathcal{C}_p(d\gamma_n).$$

Preuve. — Si $Z_n^\omega(\{x\}) = 1$ et $k < n$, $Y_k^\omega(x) = Y_k^\omega(\pi^\omega(x))$ donc :

$$E \left[\int_{\mathbb{R}} \varphi(Y_1^\omega(x), \dots, Y_n^\omega(x)) Z_n^\omega(dx) \mid \mathcal{F}_{n-1} \right] \\ = E \left[\int_{\mathbb{R}} \varphi(Y_1^\omega(y), \dots, Y_{n-1}^\omega(y), \gamma(x, y)) Z_n^\omega(dx) \mid \mathcal{F}_{n-1} \right]$$

où $y = \pi^\omega(x)$ et $\gamma(x, y) = (x - y, \omega_{n, i(y, Z_{n-1}^\omega)})$ d'où :

$$E \left[\int_{\mathbb{R}} \dots \mid \mathcal{F}_{n-1} \right] \\ = E \left[\int_{\mathbb{R}^2} \varphi(Y_1^\omega(y), \dots, Y_{n-1}^\omega(y), \gamma(x, y)) \tau_y \omega_{n, i(y, Z_{n-1}^\omega)}(dx) Z_{n-1}^\omega(dy) \mid \mathcal{F}_{n-1} \right] \\ = \int_{\mathbb{R} \times M \times \mathbb{R}} \varphi(Y_1^\omega(y), \dots, Y_{n-1}^\omega(y), (z, \Phi)) \Phi(dz) P(d\Phi) Z_{n-1}^\omega(dy),$$

d'où le résultat par récurrence.

PROPOSITION. — $\mathcal{M}_1(C)$ ensemble des probabilités sur C est muni de la topologie de la convergence étroite. On note $\mathcal{C}_p^\theta(dx, d\Phi) = \exp[\theta x - L(\theta)] \mathcal{C}_p(dx, d\Phi)$;

(1) si $A \subset \mathcal{M}_1(\mathbb{C})$:

$$\overline{\lim}_n n^{-1} \log \mathcal{P}_n^\omega \left[n^{-1} \sum_{i=1}^n \delta_{\gamma_i} \in A \right] \leq -\inf \{K(\mathcal{C}_p^0, \mathcal{D}), \mathcal{D} \in \overline{A}\},$$

p. s. sur S, où K est l'information de Kullback;

(2) si :

$$e_0 < a < L'(\theta_M), \quad h_0(a) < \log m \quad \text{et} \quad \theta = h'_0(a).$$

pour tout V voisinage de \mathcal{C}_p^0 et tout $\varepsilon > 0$ on a p. s. sur S, pour n assez grand :

$$\mathcal{P}_n^\omega \left(n^{-1} \sum_{i=1}^n \delta_{\gamma_i} \in V \mid n^{-1} \sum_{i=1}^n \gamma_i^1 \in [a, a + \varepsilon) \right) \geq 1 - e^{-Kn} \quad \text{avec} \quad K > 0.$$

Preuve. – (1) Soit $G_n = \text{card} \{ \gamma \in \Gamma_n^\omega : n^{-1} \sum \delta_{\gamma_i} \in \overline{A} \}$. Le lemme entraîne $m^{-n} EG_n = \int 1_{\overline{A}}(\delta_\gamma) (\mathcal{C}_p^0)^{*n}(d\gamma)$. C est polonais car fermé dans $\mathbb{R} \times M$ polonais. Le théorème de Chernov pour les lois empiriques [3] donne :

$$\overline{\lim}_n n^{-1} \log \int 1_{\overline{A}}(\delta_\gamma) (\mathcal{C}_p^0)^{*n}(d\gamma) \leq -\inf \{K(\mathcal{C}_p^0, \mathcal{D}), \mathcal{D} \in \overline{A}\}.$$

Or $\log \mathcal{P}_n^\omega [n^{-1} \sum \delta_{\gamma_i} \in \overline{A}] = \log G_n - \log Z_n^\omega(\mathbb{R})$.

L'hypothèse (1) entraîne $n^{-1} \log Z_n(\mathbb{R}) \rightarrow m$ p. s. sur S [1]. On conclut avec Tchébitchev et Borel-Cantelli.

(2) Étudions $n^{-1} \log EH_n^\omega$ où $H_n^\omega = \text{card} \{ \gamma \in \Gamma_n : n^{-1} \sum \delta_{\gamma_i} \in V^c, n^{-1} \sum \gamma_i^1 \in [a, a + \varepsilon) \}$. Prenons V intersection finie de demi-espaces. En utilisant la troncature proposée dans [4] pour les fonctions linéaires de lois empiriques on trouve :

$$\overline{\lim}_n [-\log m + n^{-1} \log EH_n^\omega] \leq -\inf \{K(\mathcal{C}_p^0, \mathcal{D}), \mathcal{D} \in B\},$$

où :

$$B = \left\{ \mathcal{D} \in V^c : \exists k > 0, \mathcal{D}([-k, +k] \times M) = 1 \text{ et } \int_{\mathbb{R} \times M} x \mathcal{D}(dx, d\Phi) \in [a, a + \varepsilon) \right\},$$

si :

$$\mathcal{D} \in B, \quad K(\mathcal{C}_p^0, \mathcal{D}) = K(\mathcal{C}_p^0, \mathcal{D}) + \theta \int_{\mathbb{R} \times M} x \mathcal{D}(dx, d\Phi) - L(\theta) + \log m \geq K(\mathcal{C}_p^0, \mathcal{D}) + h_0(a),$$

B est inclus dans un fermé ne contenant pas \mathcal{C}_p^0 donc :

$$\overline{\lim}_n n^{-1} \log EH_n^\omega < -h_0(a).$$

Tchebitchev et Borel-Cantelli entraînent $\overline{\lim} n^{-1} \log H_n^\theta < -h_0(a)$ p. s. sur S, ce qui avec (4) permet de conclure.

III. INTERPRÉTATION. — Les individus situés près de na au temps n ont eu en moyenne une généalogie construite à l'aide de \mathcal{C}_p^θ avec $\theta = h_0'(a)$. Par rapport à \mathcal{C}_p^0 obtenue en moyennant sans restriction, ou en restant près de ne_0 , il y a modification exponentielle. Deux décompositions de \mathcal{C}_p^θ donnent deux constructions d'une n -généalogie typique.

(1) Soit $\Phi^\theta(dx) = \exp[\theta x - \log \int_{\mathbb{R}} e^{\theta t} \Phi(dt)] \Phi(dx)$ lorsque ceci a un sens et :

$$P^\theta(d\Phi) = \left[\int_{\mathbb{R}} \exp(\theta t - L(\theta)) \Phi(dt) \right] P(d\Phi),$$

probabilité sur M portée par \dot{M} . Alors $\mathcal{C}_p^\theta(dx, d\Phi) = \Phi^\theta(dx) P^\theta(d\Phi)$.

On tire une mesure ponctuelle φ_1 suivant P^θ puis un point x_1 dans le support de φ_1 suivant la loi φ_1^θ . x_1 devient l'origine de φ_2 tirée suivant P^θ et on choisit x_2 suivant $\tau_{x_1} \varphi_2^\theta \dots$. La n -généalogie est $(x_1, \varphi_1) \dots (x_n, \varphi_n)$.

(2) Soit $\mu^\theta(dx) = e^{\theta x - L(\theta)} \mu(dx)$. On sait que $\mathcal{C}_p(dx, d\Phi) = \mu(dx) P_x(d\Phi)$ où P_x est la distribution de Palm [5]. On tire un point x_1 suivant μ^θ puis φ_1 une mesure ponctuelle suivant P_{x_1} , c'est-à-dire suivant P sachant qu'un point est déjà fixé en x_1 . Puis on tire x_2 suivant μ^θ , φ_2 suivant P_{x_2} qu'on translate de $x_1 \dots$

Pour $\theta=0$ on retrouve les constructions utilisées en [6].

(¹) Rem. si $a < e_0$, on prend $\varepsilon < 0$.

(*) Remise le 25 mai 1981.

[1] ATHREYA-NEY, *Branching Processes*, Springer Verlag, 1972.

[2] BIGGINS, *J. Appl. Prob.*, 14, 1977, p. 630-636.

[3] Séminaire Orsay, 1977-1978, *Grandes déviations et applications statistiques (Astérisque, n° 68)*.

[4] GROENBOOM, OOSTERHOFF et RUYMGAART, *Ann. Prob.*, 7, 1979, p. 553-586.

[5] KALLENBERG, *Random Measures*, Academic Press, 1976.

[6] KALLENBERG, *Math. Nachr.*, 77, 1977, p. 7-43.

[7] MATTHES, KERSTAN et MECKE, *Infinitely Divisible Point Processes*, Wiley, 1978.

E.R.A., C.N.R.S. 532, Université Paris-Sud,
Mathématiques, Bât. 425, 91405 Orsay Cedex.

Note supplémentaire :

Depuis la parution de cet article, O. Nerman nous a fait connaître les travaux de l'école suédoise, dont certains ont rapport à l'échantillonnage sur une population (supercritique) d'un processus temporel.

JAGERS (P.), NERMAN (O.), *The growth and composition of branching populations*, *Adv. Appl. Prob.*, 1984, t.16, p. 221-259.

NERMAN (O.), JAGERS (P.), *The stable doubly infinite pedigree process of supercritical branching populations*. *Zeits. für Wahr.* 1984, t. 65, p. 445-460.

F. Probabilités de présence dans un processus de branchement
spatial markovien.

PROBABILITES DE PRESENCE DANS UN PROCESSUS DE BRANCHEMENT
 SPATIAL MARKOVIEN

Alain ROUAULT
 Unité Associée 743 - A.D.4.
 Mathématique (Bât. 425)
 91405 - ORSAY Cedex.

Résumé : - Pour un processus de Galton-Watson $(Z_n)_{n \in \mathbb{N}}$ sous-critique, $n^{-1} \text{Log } P(Z_n > 0)$ et $n^{-1} \text{Log } EZ_n$ ont même limite $\text{Log } m$. Nous avons étendu ce résultat [8] aux processus de branchement spatiaux homogènes au voisinage d'un point en zone dite sous-critique. Nous étudions ici le cas markovien dans l'asymptotique naturelle des grandes déviations et montrons que, sous une hypothèse de sous-criticité, on a un résultat similaire pour la probabilité de présence au "voisinage" de x , le rôle de $-\text{Log } m$ étant jouée par l'inf des intégrales d'action de chemins aboutissant en x .

Abstract : - For a subcritical Galton-Watson process $(Z_n)_{n \in \mathbb{N}}$, $n^{-1} \text{Log } P(Z_n > 0)$ and $n^{-1} \text{Log } EZ_n$ have the same limit $\text{Log } m$. We extended this result [8] to branching random walks in the neighbourhood of a point in an area called subcritical. Here we study markovian case in the natural framework of large deviations and show, under a subcritical hypothesis, similar result for the presence probability in the neighbourhood of x , with $-\text{Log } m$ replaced by the infimum of action integrals of paths reaching x .

PROBABILITES DE PRESENCE DANS UN PROCESSUS DE BRANCHEMENT
SPATIAL MARKOVIEN

Alain ROUAULT

INTRODUCTION :

Un processus de Galton-Watson $(Z_n)_{n \geq 0}$ de moyenne m , avec $Z_0=1$, vérifie

$$i) E Z_n = m^n$$

S'il est sous critique ($m < 1$) il vérifie de plus :

$$ii) \lim_n Z_n = 0 \quad \text{p.s.}$$

$$iii) \lim_n \frac{1}{n} \text{Log } P(Z_n > 0) \geq \text{Log } m \text{ dès que } E Z_1 \text{ Log } Z_1 < \infty \quad [5]$$

i) et iii) ont comme conséquence immédiate (inégalité de Markov)

$$iv) \lim_n \frac{1}{n} \text{Log } P(Z_n > 0) = \text{Log } m$$

L'évolution d'un processus de branchement spatial homogène Z_n sur \mathbb{R} (mesure de comptage sur \mathbb{R}) est décrite dans [2], [3]. Si $\mu = E Z_1$ est l'espérance du processus de reproduction on a facilement $E Z_n = \mu^{*n}$ et donc, si h est la transformée de Cramer de μ (duale de Legendre de la Log-Laplace de μ , supposée définie partout [1], [4]), pour tout (a, δ) de $\mathbb{R} \times \mathbb{R}_*^+$ si Δ_a^δ désigne $[a, a+\delta)$ ou $(a-\delta, a]$ suivant le signe de $h'(a)$, on a notamment :

$$i') \lim_n \frac{1}{n} \text{Log } E Z_n (n \Delta_a^\delta) = -h(a)$$

La fonction $-h$ joue le rôle de $\text{Log } m$ et si $h(a) > 0$ on a [2]

$$ii') \lim_n Z_n (n \Delta_a^\delta) = 0 \quad \text{p.s.}$$

Dans [8] nous avons appelé zone sous critique la région $h > 0$ et montré l'analogie

de iii) sous une hypothèse de moment d'ordre 2 pour le nombre d'enfants par individu. On se propose ici, dans le cas spatial markovien non homogène de montrer, sous une hypothèse $(k \log k)$ le résultat type iii) correspondant, conjecturé en [11] et énoncé en [9]. Le résultat type i), beaucoup plus simple est traité comme cas particulier en [10] .

On se donne une famille indexée par $\mathbb{R}^+ \times \mathbb{R}$ (temps, espace) de lois de probabilité sur M , espace des mesures ponctuelles finies sur \mathbb{R} . L'élément générique est $P_{t,x}(d\nu)$, ayant une espérance notée $\mu_{t,x}$, mesure de Radon positive sur \mathbb{R} . $P_{t,x}$ décrit, à l'échelle 1, la loi des positions des enfants d'un individu situé en x à l'instant t . Si besoin est, on particularisera $P_{t,x}$ en supposant qu'il est découplé, c'est-à-dire construit à partir d'un nombre aléatoire N (de fonction génératrice $g_{t,x}$) de copies de v.a. indépendantes tirées suivant une loi $\mu_{t,x}^0$. Le cas homogène correspond à $P_{t,x} \equiv P$ pour tout t,x .

On construit une famille indexée par ε tendant vers 0 de processus de branchement spatiaux. Pour ce faire, on se donne $x_0 \in \mathbb{R}$, position d'un ancêtre (individu de la génération 0). Il vit un temps ε puis est remplacé par des enfants dont les positions forment un élément (aléatoire) de M obtenu de la manière suivante : on tire un élément de M suivant P_{ε,x_0} , on lui applique une homothétie de rapport ε puis une translation de vecteur x_0 . Les enfants de l'ancêtre forment la première génération et sont datés de l'instant ε , de sorte qu'un tel individu, situé en x , sera à l'instant 2ε remplacé par des enfants dont les positions formeront un élément de M obtenu par tirage suivant $P_{2\varepsilon,x}$, ε -homothétie et x -translation. On itère la procédure et on s'intéresse pour tout t à l'élément de M formé par les individus en vie à t (donc de génération $\lceil \frac{t}{\varepsilon} \rceil$).

L'asymptotique est celle des grandes déviations : les reproductions sont de plus en plus fréquentes (instants $k\varepsilon$, $k \in \mathbb{N}$) et les sauts de plus en plus petits

(rapport ε). L'outil fondamental est l'intégrale d'action ([1], [4], [12] : pour les marches aléatoires) définie sur $A[t_0, t_1]$, ensemble des fonctions absolument continues de $[t_1, t_2]$ dans \mathbb{R} par

$$I(t_1, t_2, \varphi) = \int_{t_1}^{t_2} h(s, \varphi(s), \dot{\varphi}(s)) ds$$

où $h(t, x, \cdot)$ est la transformée de Cramer [13] de la mesure $\mu_{t, x}$.

Un chemin doit être vu comme la limite (en ε) d'une ligne généalogique constituée par les positions des ancêtres successifs d'un individu. Dans le cas homogène, pour des raisons de convexité, ces chemins sont rectilignes. La présence éventuelle de population près de x à l'instant t s'étudie à partir de l'existence de lignes généalogiques dans des "tubes" autour de chemins conduisant de

x_0 à x . On montre que si $\inf_{\substack{\varphi(0)=x_0 \\ \varphi(t)=x}} I(0, t, \varphi)$ est atteint pour un φ de classe

C^1 vérifiant "grosso-modo" $\inf_{0 \leq s \leq t} h(s, \varphi(s), \dot{\varphi}(s)) > 0$ (notion de sous-criticité)

alors l' ε -logarithme de la probabilité, partant de x_0 de présence près de x à l'instant t est, à la limite minorée par (en fait, tend vers, grâce à [10])

$$- \inf_{\substack{\varphi(0)=x_0 \\ \varphi(t)=x}} I(0, t, \varphi).$$

Ayant besoin de lignes généalogiques, on construit au § I un espace de probabilité sur les arbres spatiaux aléatoires dans l'esprit de [7]. § II sert à présenter les outils de grandes déviations utilisés par la suite. Au § III on établit une minoration de probabilité de présence dite exacte (non-asymptotique). Enfin au § IV on établit les théorèmes limites, d'abord relativement à un tube, puis avec optimisation.

§ I - ARBRES SPATIAUX ALEATOIRES.

Introduction :

Pour l'étude de notre modèle, il nous faut construire un processus de branchement spatial assez général. Le processus le plus simple, que l'on peut appeler homogène est obtenu en conjuguant deux phénomènes aléatoires simples : un processus de Galton-Watson et une marche aléatoire. Chaque individu d'un Galton-Watson est muni d'une position aléatoire dans \mathbb{R} et les variations de position d'une génération à l'autre sont obtenues par des copies indépendantes d'une même loi. Le processus général consiste à :

- 1) se donner les positions des descendants immédiats d'un individu globalement, c'est-à-dire sous la forme d'une mesure ponctuelle ;
- 2) imposer un caractère markovien : la loi de cette mesure ponctuelle ne dépend que de la position de l'individu considéré (et de l'instant) ;
- 3) garder la propriété de branchement : indépendance des descendance d'individus différents appartenant à la même génération.

Dans plusieurs études nous avons eu besoin de renseignements précis sur l'histoire d'un individu, c'est-à-dire les positions de ses différents ascendants. Un formalisme adéquat est fourni par une extension au cas markovien des arbres marqués de Neveu [7].

Mesures ponctuelles :

Soit \bar{M}_k l'ensemble des mesures de Radon positives sur \mathbb{R}^k , $B(\bar{M}_k)$ la tribu borélienne pour la topologie vague. L'ensemble M_k des mesures ponctuelles finies (sommées finies de masses de Dirac) appartient à $B(\bar{M}_k)$ et est muni de la tribu trace notée $B(M_k)$. On note M pour M_1 .

Si a désigne un point extérieur à \mathbb{R} et δ_x la masse de Dirac en x , on définit facilement une indexation mesurable, c'est-à-dire une application mesurable

φ de $\mathbf{N}_* \times M$ dans $\mathbb{R} \cup \partial$ telle que :

$$\mu = 0 \implies \varphi(k, \mu) = \partial \quad \text{pour tout } k$$

$$\mu \neq 0 \implies \varphi(k, \mu) = \partial \quad \text{pour } k > \mu(1)$$

$$\text{et } \mu = \sum_{k=1}^{\mu(1)} \delta_{\varphi(k, \mu)}$$

Autrement dit, à l'extension naturelle près, $\varphi(k, \mu)$ désigne le k -ième point de μ . On notera $\varphi_k(\mu)$ pour $\varphi(k, \mu)$. Voir par exemple [15].

Espace probabilisable des arbres marqués :

L'outil présenté dans ce paragraphe est dû à Neveu [7]. Nous en énonçons succinctement les grandes lignes, avec de légères modifications de notations, et l'appliquons à notre modèle.

U désigne l'espace des suites finies $u = j_1 \dots j_n$ d'éléments de \mathbf{N}_* . La suite vide ϕ est dans U . La longueur de $u \in U$ se note $|u|$. Pour tout $n \in \mathbf{N}$, $\{u \in U : |u| = n\}$ coïncide avec \mathbf{N}_*^n et donc

$$U = \sum_{n \geq 0} \mathbf{N}_*^n$$

On note f la fonction de U dans U définie par :

$$f(\phi) = \phi, \quad f(j_1 \dots j_n) = j_1 \dots j_{n-1} \quad \text{pour tout } n \geq 1$$

La concaténée de u et v de U se note uv .

Un arbre est par définition un sous-ensemble $\tilde{\omega}$ de U , vérifiant

$$* \quad \text{pour tout } n \quad \tilde{\omega} \cap \mathbf{N}_*^n \text{ est fini}$$

$$** \quad f(\tilde{\omega}) \subset \tilde{\omega}$$

Remarquons que $**$ entraîne $\phi \in \tilde{\omega}$. Les éléments de $\tilde{\omega}$ sont ses noeuds, $\tilde{\Omega}$ est l'espace des arbres et pour $u \in U$ $\tilde{\Omega}_u = \{\tilde{\omega} : u \in \tilde{\omega}\}$. On note \tilde{I} la tribu engendrée par les $\tilde{\Omega}_u$. Pour $u \in U$, on note $\tilde{\theta}_u$ l'application mesurable de $\tilde{\Omega}_u$ dans $\tilde{\Omega}$ définie par :

$$\tilde{\Theta}_u(\tilde{\omega}) = \{v : v \in U \text{ et } uv \in \tilde{\omega}\}$$

(On a translaté l'arbre à son noeud u).

L'espace mesurable $M, \mathcal{B}(M)$ est l'espace des marques. Un arbre marqué sera défini par la donnée d'un arbre $\tilde{\omega} \in \tilde{\Omega}$ et d'une marque $v_u \in M$ pour chaque noeud $u \in \tilde{\omega}$ soit

$$\omega = (\tilde{\omega}; (v_u, u \in \tilde{\omega}))$$

En fait nous marquons chaque individu avec la mesure ponctuelle représentant les positions de ses enfants.

Ω désigne l'espace des arbres marqués et $p : \Omega \rightarrow \tilde{\Omega}$ la projection canonique. Si $u \in U$, v_u est une application de $\Omega_u = p^{-1}(\tilde{\Omega}_u)$ dans M . Ω est munie de la tribu \mathcal{I} engendrée par $p^{-1}(\mathcal{I})$ et les applications $v_u : \Omega_u \rightarrow M$ ($u \in U$). On note pour tout n , F_n la tribu engendrée par les Ω_u tels que $|u| \leq n$ et par les v_u tels que $|u| < n$.

Ω est muni des translations $\Theta_u : \Omega_u \rightarrow \Omega$ définies par

$$\Theta_u(\omega) = [\tilde{\Theta}_u(\tilde{\omega}), (v_{uv}, v \in \tilde{\Theta}_u(\tilde{\omega}))] \text{ si } \omega = (\tilde{\omega}, (v_u, u \in \tilde{\omega}))$$

On a

$$p \circ \Theta_u = \tilde{\Theta}_u \circ p \quad \text{sur } \Omega_u = p^{-1}(\tilde{\Omega}_u)$$

Si on note $v = v_\phi$ on a pour tout u de U

$$v_u = v \circ \Theta_u \quad \text{sur } \Omega_u \tag{1.1}$$

si $u \in U$ avec $|u| \geq 1$, $u = j_1 \dots j_n$ on note

$$\sigma(u, \omega) = \varphi_{j_n}(v_{j_1 \dots j_{n-1}}) \quad \text{pour } \omega \in \Omega_u$$

($\sigma(u, \omega)$ donne la position de l'individu identifié au noeud u).

L'application $\sigma(u) = \sigma(u, \cdot)$ de Ω_u dans $\bar{\mathbb{R}} = \bar{\mathbb{R}} \cup \{\partial\}$ est $\Omega_u \cap F_n$ mesurable.

Dans les mêmes conditions on note

$$s(u) = (\sigma \circ f^{n-1}(u), \sigma \circ f^{n-2}(u), \dots, \sigma \circ f(u), \sigma(u))$$

application de Ω_u dans \mathbb{R}^n , $\Omega_u \cap F_n$ mesurable.

Dans le but d'étudier la $n^{\text{ième}}$ génération, $n \geq 1$, on définit

$$z_n(\omega) = p(\omega) \cap \mathbf{N}_*^n$$

\tilde{I} est engendrée par les $(z_n, n \in \mathbf{N})$. On note \tilde{F}_n la tribu engendrée par (z_0, \dots, z_n) .

Construction de l'espace probabilisé.

Nous probabilisons Ω en adaptant la construction de [7] au cas markovien : celle-ci consiste à prendre un espace canonique Ω^* plus gros et de structure plus simple. En fait $\Omega^* = M^U$ convient mais la structure markovienne exige quelques vérifications.

Proposition 1 :

Pour toute famille mesurable $P_{i,x}$, $(i,x) \in \mathbf{N} \times \mathbb{R}$ de probabilités sur $(M, \mathcal{B}(M))$ il existe une (unique) famille mesurable $P_{i,x}$, $(i,x) \in \mathbf{N} \times \mathbb{R}$ de probabilités sur (Ω, \mathcal{I}) telle que pour tout (i,x) de $\mathbf{N} \times \mathbb{R}$:

- 1) la loi de v sous $P_{i,x}$ est $P_{i,x}$
- 2) pour tout u de U on a $P_{i,x}(\Omega_u \cap \{\sigma(u) = \partial\}) = 0$
- 3) pour tout $n > 0$, conditionnellement à F_n sur $(\Omega, \mathcal{I}, P_{i,x})$ les v.a.

Θ_u ($u \in z_n$) sont indépendantes et de lois respectives $P_{\bar{\sigma}(u)}$ où $\bar{\sigma}$ fonction de U dans $\mathbf{N} \times \mathbb{R}$ est définie par

$$\bar{\sigma}(\phi) = (i,x) \quad \bar{\sigma}(u) = (i+|u|, \sigma(u)) \quad \text{si } |u| \geq 1$$

Ceci signifie que pour tout choix des fonctions mesurables positives g_u de Ω dans \mathbb{R} on a :

$$E_{i,x}^{F_n} \left[\prod_{u \in z_n} g_u \circ \Theta_u \right] = \prod_{u \in z_n} E_{\bar{\sigma}(u)} (g_u)$$

Démonstration succincte :

Soit $\Omega^* = M^U$ muni des applications coordonnées v_u^* de Ω^* dans M , de la tribu canonique I^* et pour tout $n \geq 1$ de F_n^* engendrée par tous les v_u^* pour

$|u| \leq n-1$. Pour tout (i,x) on construit sur (Ω^*, I^*) une probabilité par prolongement (Kolmogorov) d'un système projectif, indexé par l'ensemble des arbres finis, muni de la relation d'ordre naturelle (inclusion), par récurrence sur la hauteur des arbres. En passant d'une hauteur à la suivante on impose évidemment les propriétés requises. On définit la position $\sigma^*(u)$ de $u = j_1 \dots j_n \in U$ par $\sigma^*(u) = \varphi_{j_n}^*(v_{j_1 \dots j_{n-1}}^*)$ et $\sigma^*(\phi) = x$. L'espace (Ω^*, I^*) est trop gros car il introduit des positions cimetières (ou individus fictifs). On transporte cette structure sur Ω grâce à l'application ψ de Ω^* dans Ω définie par :

$$\omega^* \in \Omega^* \quad \psi(\omega^*) = \{(u, v_u^*(\omega^*)) : \forall k \leq |u| \quad \sigma^*(f^k u) \neq \partial\}$$

On vérifie que $p(\psi(\omega^*))$ est un arbre, que ψ est mesurable, que pour tout n on a $\psi^{-1}(F_n) \subset F_n^*$ et que $\theta_u \circ \psi = \psi \circ \theta_u^*$ sur $\psi^{-1}(\Omega_u)$. On appelle $P_{i,x}$ l'image de $P_{i,x}^*$ par ψ . Les propriétés 1,2,3 sont alors des conséquences de la construction (voir [7] dans le cas homogène).

Conséquences : Mesures de comptage.

Pour tout (i,x) , $P_{i,x}$ ne charge que $\psi(\Omega^*)$ sur lequel on définit, (les sommes vides valant 0 par convention), pour tout $n \geq 1$

$$\zeta_n(\omega) = \sum_{u \in Z_n(\omega)} \delta_{\sigma(u)} \quad , \quad Z_n(\omega) = \sum_{u \in Z_n(\omega)} \delta_{s(u)}$$

Par construction ζ_n (resp. Z_n) est une application F_n -mesurable de Ω dans $M, \mathcal{B}(M)$ (resp. $M_n, \mathcal{B}(M_n)$). On note $\zeta_n(\omega, g)$ (resp. $Z_n(\omega, g)$) ou $\zeta_n(g)$ (resp. $Z_n(g)$) la valeur de ζ_n (resp. Z_n) appliquée à la fonction g .

Si n est un entier positif fixé, on note U^n l'ensemble des suites finies de longueur au plus n de nombres réels, y compris la suite vide. On peut écrire si \mathbb{R}^0 désigne $\{\phi\}$, $U^n = \sum_0^n \mathbb{R}^k$. On note $|\phi| = 0$ et si $0 < k \leq n$ et $\ell \in \mathbb{R}^k$, $|\ell| = k$. La concaténation, notée $[\cdot, \cdot]$, se définit de manière naturelle par :

$$\begin{aligned} \text{si } \ell \in U^n & \quad [\phi, \ell] = [\ell, \phi] = \ell \\ \text{si } \ell^1, \ell^2 \in U^n & \quad , \quad |\ell^1| = k_1 \quad , \quad |\ell^2| = k_2 \quad , \quad k_1 + k_2 \leq n \\ & \quad [\ell^1, \ell^2] = (\ell_1^1, \dots, \ell_{k_1}^1, \ell_1^2, \dots, \ell_{k_2}^2) \end{aligned}$$

Si γ est une fonction mesurable de \mathbb{R}^n dans \mathbb{R} , on définit, pour chaque ℓ vérifiant $|\ell| \leq n-1$, la fonction γ_ℓ de $\mathbb{R}^{n-|\ell|}$ dans \mathbb{R} par

$$\gamma_\ell(\ell') = \gamma([\ell, \ell'])$$

D'après la définition des opérateurs θ de translation on a, pour $0 < k \leq n$ et $\omega \in \Omega$:

$$Z_n(\omega, \gamma) = \sum_{u \in Z_{n-k}(\omega)} Z_k(\theta_u(\omega), \gamma_S(u)) \quad (1.3)$$

et compte tenu de la proposition 1, à k, γ fixés les variables aléatoires

$\omega \rightarrow Z_k(\theta_u(\omega), \gamma_S(u))$, $\omega \in Z_{n-k}$ sont pour tout (i, x_0) de $\mathbb{N} \times \mathbb{R}$

F_{i, x_0}^{n-k} - indépendantes. Pour $k=1$ on a

$$Z_n(\gamma) = \sum_{u \in Z_{n-1}} \nu_u(\gamma_S(u)) \quad (1.4)$$

On supposera en fait que pour tout $k, x, P_{k,x}$ a une espérance notée $Q(k, x, \cdot)$, mesure de Radon sur \mathbb{R} , c'est-à-dire que pour toute fonction de $K(\mathbb{R})$ (continue à support compact), si ν est l'élément générique de M on a

$$E_{k,x} \nu(f) = \iint_{\mathbb{R} \times M} f(y) \nu(dy) P_{k,x}(\nu) = Q(k, x, f)$$

On précise maintenant les propriétés demandées à Q .

§ II - NOYAUX ET CRAMER-CHERNOV.

Ce paragraphe qui rassemble les propriétés requises par les mesures de transition - ou noyaux - est directement inspiré de Ventseï [4] [12], étendu au cas de masses différentes de 1, en liaison avec l'usage ultérieur pour les arbres spatiaux aléatoires.

1. Présentation :

Soit Q une fonction de $\mathbf{N} \times \mathbf{R} \times \mathcal{B}(\mathbf{R})$ dans \mathbf{R}^+ dite noyau, vérifiant

- a) $\forall (i,A) \in \mathbf{N} \times \mathcal{B}(\mathbf{R}) \quad x \mapsto Q(i,x,A)$ est mesurable (Borel)
 b) $\forall (i,x) \in \mathbf{N} \times \mathbf{R} \quad A \mapsto Q(i,x,A)$ est une mesure finie sur $\mathbf{R}, \mathcal{B}(\mathbf{R})$

Q est dit markovien (resp. sous-markovien) si pour tout (i,x) de $\mathbf{N} \times \mathbf{R}$ on a $Q(i,x,\mathbf{R}) = 1$ (resp. ≤ 1). Pour tout (k,i,x) de $\mathbf{N}_* \times \mathbf{N} \times \mathbf{R}$ on définit sur \mathbf{R}^k la mesure $Q_k(i,x, \cdot)$ par la formule ($f \in \mathcal{K}(\mathbf{R}^k)$)

$$Q_k(i,x,f) = \int_{\mathbf{R}^k} f(y_1, \dots, y_k) Q(i,x, dy_1) Q(i+1, y_1, dy_2) \dots Q(i+k-1, y_{k-1}, dy_k) .$$

Remarque 1 :

Si $Q(i,x, \cdot)$ est l'espérance de $P_{i,x}$ du § I, on déduit de la proposition 1 et de (1.4)

$$E_{i,x}^{F_{n-1}} Z_n(\gamma) = \sum_{u \in Z_{n-1}} Q(\bar{\sigma}(u), \gamma_S(u)) \quad (2.1)$$

et par récurrence

$$E_{i,x} Z_n(\gamma) = Q_n(i,x,\gamma) \quad (2.2)$$

On suppose que la fonction ψ de $\mathbf{N} \times \mathbf{R}$ dans \mathbf{R} définie par

$$\psi(i,x,\theta) = \int_{\mathbf{R}} \exp \theta(\xi-x) Q(i,x,d\xi)$$

est finie partout et on note

$$G(i,x,\theta) = \text{Log } \Psi(i,x,\theta) \quad (2.3)$$

Pour tout (i,x) on sait ([1], [4]) que la fonction $G(i,x, \cdot)$ est convexe et

de classe C^∞ et on suppose que l'enveloppe convexe du support de $Q(i,x,\cdot)$ est \mathbb{R} , ce qui implique que sa transformée de Cramer [13], $H(i,x,\cdot)$ définie par

$$H(i,x,v) = \sup_{\theta} \theta v - G(i,x,\theta)$$

est finie partout et de classe C^∞ (en v).

Remarque 2 :

S'il existe $M \geq 0$ tel que

$$\inf_{i,x,v} H(i,x,v) \geq -M \quad (2.5)$$

alors $e^{-M}Q$ est un noyau sous-markovien. En effet (2.3) entraîne

$\log Q(i,x,1) = G(i,x,0)$ et par dualité $G(i,x,0) = \sup_v -H(i,x,v) \leq M$. On

supposera désormais (2.5) vérifiée.

Pour toute fonction θ de $\mathbb{R}^+ \times \mathbb{R}$, on définit pour $(i,x,\xi) \in \mathbb{N} \times \mathbb{R}^2$

$$\rho_{i,x}^\theta(\xi) = \exp \{ \theta(i,x) (\xi-x) - G(i,x,\theta(i,x)) \} \quad (2.6)$$

et le nouveau noyau Q^θ par :

$$Q^\theta(i,x,d\xi) = \rho_{i,x}^\theta(\xi) Q(i,x,d\xi) \quad (2.7)$$

Q^θ est toujours markovien et sa log-Laplace G^θ est fournie par la formule :

$$G^\theta(i,x,\eta) = G(i,x,\theta(i,x)+\eta) - G(i,x,\theta(i,x)) \quad (2.8)$$

Sur $\mathbb{R}^{\mathbb{N}}$ on définit π_i , $i = 0,1,\dots$ la suite des applications coordonnées et pour $0 \leq i < j$, on note $G_{i,j}$ la tribu engendrée par π_i, \dots, π_j . Pour $\ell \in \mathbb{R}^j$ et $i < j$ on pose :

$$R_{i,j}^\theta(\ell) = \prod_{k=i}^{j-1} \rho_{k,\ell_k}^\theta(\ell_{k+1}) \quad (2.9)$$

$R_{i,j}^\theta$ est $G_{i,j}$ mesurable, on note $\tilde{R}_{i,j}^\theta$ sa projection sur $\mathbb{R}^{\{i,\dots,j\}}$. Particularisons θ . Soit $\lambda \in \mathbb{R}^{\mathbb{N}}$ et θ la fonction de $(\mathbb{N} \times \mathbb{R})$ dans \mathbb{R} définie par

$$(k,y) \rightarrow \frac{\partial H}{\partial v}(k,y, \lambda_{k+1} - \lambda_k) \quad (2.10)$$

(Nous forçons, à l'instant k , le déplacement sous-jacent à avoir la vitesse moyenne $\lambda_{k+1}-\lambda_k$).

Notons pour $\ell \in \mathbb{R}^{\mathbb{N}}$.

$$U_i^\lambda(\ell) = \ell_i - \lambda_i \quad i \in \mathbb{N} \quad (2.11)$$

$$V^\lambda(i, j, \lambda) = \theta(i, \lambda_i) \{ \ell_{i+1} - \ell_i - \lambda_{i+1} + \lambda_i \} + \sum_{k=i}^{j-1} \theta(k, \lambda_k) \{ \ell_{k+1} - \ell_k - \lambda_{k+1} + \lambda_k \}$$

U_i^λ (resp. $V^\lambda(i, j)$) est $G_{i, i}$ (resp. $G_{i, j}$) mesurable. La dualité relie de manière simple (2.9) et (2.11) par :

$$\text{Log } R_{i, j}^\theta(\ell) = V^\lambda(i, j, \ell) + \sum_{k=i}^{j-1} H(k, \ell_k, \lambda_{k+1} - \lambda_k) \quad (2.12)$$

Q^θ étant markovien, on note encore π la chaîne canonique associée. (2.8) entraîne, avec des notations évidentes,

$$E_{i, x}^\theta (\pi_{i+1} - \pi_i) = \frac{\partial G}{\partial \theta} (i, x, \theta(i, x))$$

et donc sous Q_{i, λ_i}^θ les 2 suites $(0, (U_j^\lambda)_{j>i})$ et $(0, (V^\lambda(i, j))_{j>i})$ sont des martingales adaptées à $(G_{i, j})_{j \geq i}$.

Nous allons maintenant introduire une notion de tube généralisé sur lequel la dérivée de Randon-Nikodym R^θ sera presque constante, ce qui permettra des évaluations de la mesure du tube, grâce à l'inégalité de Kolmogorov.

2. Evaluations sur les tubes.

Pour $i < j$ et $(u, v, \lambda) \in (\mathbb{R}_*^+)^2 \times \mathbb{R}^{\mathbb{N}}$

$$\Gamma^\lambda(i, j, u, v) = \prod_{k=i}^j \{ |U_k^\lambda| \leq u, |V^\lambda(i, k)| \leq v \} \quad (2.14)$$

qui appartient à $G_{i, j}$. On note $\tilde{\Gamma}^\lambda(i, j, u, v)$ la projection de

$\Gamma^\lambda(i, j, u, v) \cap \{ \ell : \ell_j = \lambda_j \}$ sur $\mathbb{R}^{\{i+1, \dots, j\}}$. On montre (par ex. d'après [4]) les lemmes suivants.

Lemme 1 :

Si pour $i < j$ et $\lambda \in \mathbb{R}^{\mathbb{N}}$, $u \in \mathbb{R}_*^+$ on pose

$$I_1(i, j, \lambda) = \sum_{k=i}^{j-1} H(k, \lambda_k, \lambda_{k+1} - \lambda_k) \quad (2.15)$$

et

$$\Delta(i,u) = \sup_{\substack{|k-k'| < i \\ |x-x'| < u \\ a \in \mathbb{R}}} \frac{H(k',x',a) - H(k,x,a)}{M+1+H(k,x,a)}$$

alors pour tout $v > 0$ les conditions $\ell \in \Gamma_\lambda(i,j,u,v)$ et $|\ell_i - \lambda_i| \leq u$ entraînent

$$|\text{Log } R_{i,j}^\theta(\ell) - I_1(i,j,\lambda)| \leq D_1(i,j,u,v,\lambda)$$

où

$$D_1(i,j,u,v,\lambda) = v + \Delta(0,u) (M+1) (j-i) + I_1(i,j,\lambda) \quad \blacksquare$$

Lemme 2 :

Avec les notations du lemme 1 on pose

$$D(i,j,\lambda,u,v) = I_1(i,j,\lambda) + \text{Log } Q_{j-i}(i,\lambda_i, \tilde{\Gamma}^\lambda(i,j,u,v))$$

$$D_2(i,j,\lambda) = \sup_{\substack{i \leq k \leq j-1 \\ x}} \frac{\partial^2 G}{\partial \theta^2}(k,x, (k,x))$$

$$D'_2(i,j,\lambda) = \sup_{\substack{i \leq k \leq j-1 \\ x}} \left\{ \frac{\partial^2 G}{\partial \theta^2}(k,x,\theta(k,x)) |\theta(k,x)|^2 \right\}$$

où λ et θ sont liés par (2.10).

On a alors

$$1) D(i,j,\lambda,u,v) \leq D_1(i,j,\lambda,u,v)$$

$$2) D(i,j,\lambda,u,v) \geq -D_1(i,j,\lambda,u,v) + \text{Log} \left\{ 1 - \frac{j D_2(i,j,\lambda)}{u^2} - \frac{j D'_2(i,j,\lambda)}{v^2} \right\} +$$

§ III - PROBABILITE DE PRESENCE - MINORATION EXACTE ET LEMMES TECHNIQUES.

Nous allons établir des résultats comparables aux lemmes 2 et 3 de [8] mais plus complexes dans la mesure où le processus n'est pas homogène et plus précis dans la mesure où nous avons une formule exacte pour p (formule 3.6) préparant l'utilisation d'une hypothèse type $(k \text{ Log } k)$ (lemme 4).

1. Définitions et notations :

On appelle processus ponctuel sur \mathbb{R} découplé, une probabilité P sur $(M, \mathcal{B}(M))$ [8] telle que P soit la loi de $\sum_{i=1}^N \delta_{\xi_i}$ où δ est la masse de Dirac, les ξ_i , $i = 1, 2, \dots$ indépendantes et de même loi μ^0 et N (à valeurs dans \mathbb{N}) indépendantes des ξ_i . P a une espérance en même temps que N et pour toute f de $K(\mathbb{R})$ on a :

$$E \nu(f) = \int_M \nu(f) P(d\nu) = (EN) \mu^0(f)$$

On se donne une famille mesurable $P_{i,x}$, $(i,x) \in \mathbb{N} \times \mathbb{R}$ de processus ponctuels découplés. Suivant la proposition 1, on construit alors la famille $P_{i,x}$, $(i,x) \in \mathbb{N} \times \mathbb{R}$ correspondante, dite processus de branchement spatial découplé. On suppose de plus que P a une espérance Q vérifiant les conditions de II.

Fixons $(n, i, x_0) \in \mathbb{N}^* \times \mathbb{N} \times \mathbb{R}$ et définissons $\bar{\tau}$ de U^n dans $\mathbb{N} \times \mathbb{R}$ par

$$\bar{\tau}(\phi) = (i, x_0)$$

$$\bar{\tau}(\ell) = (i+k, \ell_k) \quad \text{si } 0 < k \leq n \quad \ell = (\ell_1, \dots, \ell_k) \in \mathbb{R}^k \quad (3.1)$$

Soit γ mesurable à support compact S de \mathbb{R}^n dans \mathbb{R}^+ .

Si $\ell \in U^n$ on définit

$$\begin{aligned} \text{pour } |\ell| \leq n-1 \quad B(\ell) &= \{Z_{n-|\ell|}(\gamma_\ell) > 0\} \\ p(\ell) &= P_{\bar{\tau}(\ell)}(B(\ell)) \end{aligned} \quad (3.2)$$

$$\text{pour } |\ell| = n \quad p(\ell) = \gamma(\ell)$$

Si γ est une indicatrice, $p(\phi)$ est une probabilité de présence. On cherche

en fait à la comparer à l'espérance correspondante (voir introduction). On définit donc q par :

$$|\ell| \leq n-1 \Rightarrow q(\ell) = E_{\bar{\tau}(\ell)} Z_{n-|\ell|} (\gamma_\ell) \quad (3.3)$$

Comme dans [8], une autre fonction utile, r , est définie par :

$$|\ell| \leq n-1 \Rightarrow r(\ell) = \int_{\mathbb{R}} p_\ell(y) Q(\bar{\tau}(\ell), dy) \quad (3.4)$$

Si pour tout k tel que $1 \leq k \leq n$, $\pi_{1,k}$ désigne dans \mathbb{R}^n la projection sur les k premières coordonnées, alors p, q, r ont des supports inclus dans

$$\tilde{S} = \{\phi\} + \sum_{k=1}^n \pi_{1,k}(S)$$

Pour le lemme suivant, qui permet une première minoration, on note

$$N = Z_1(1) = v(1) = \text{card } z_1 \quad (\text{voir } \S I)$$

et pour tout $(k, x, s) \in \mathbb{N} \times \mathbb{R} \times [0, 1]$

$$g_{k,x}(s) = E_{k,x} s^N, \quad m_{t,x} = E_{k,x} N = g'_{k,x}(1) = Q(k, x, 1) \quad (3.5)$$

$$\beta_{k,x} = \frac{P_{k,x}(N > 0)}{E_{k,x} N}, \quad \rho_{k,x}(s) = E_{k,x} \left(N - \frac{1-s^N}{1-s} \right), \quad \rho_{k,x}(1) = 1$$

2. Lemme fondamental.

1) Pour tout ℓ de U^n tel que $|\ell| \leq n-1$ on a

$$a) \quad p(\ell) = r(\ell) \left[1 - \frac{1}{m_{\bar{\tau}(\ell)}} \rho_{\bar{\tau}(\ell)} \left(1 - \frac{r(\ell)}{m_{\bar{\tau}(\ell)}} \right) \right] \quad (3.6)$$

$$b) \quad p(\ell) \geq \beta_{\bar{\tau}(\ell)} r(\ell) \quad (3.7)$$

2) Si pour $k \leq n-1$ on note

$$\eta_k = \inf_{\substack{|\ell|=k \\ \ell \in \tilde{S}}} \left[1 - \frac{1}{m_{\bar{\tau}(\ell)}} \rho_{\bar{\tau}(\ell)} \left(1 - \frac{q(\ell)}{m_{\bar{\tau}(\ell)}} \right) \right] \quad (3.8)$$

$$\text{et si } \beta \text{ vérifie } \beta \leq \inf_{\ell \in \tilde{S}} \beta_{\bar{\tau}(\ell)} \quad (3.9)$$

alors on a

$$p(\phi) \geq q(\phi) \prod_{k=0}^{n-1} (\eta_k \vee \beta) \quad (3.10)$$

Démonstration :

L'idée consiste à conditionner par rapport à la première génération . ℓ est fixé dans la démonstration. D'après (1.3) on a :

$$B(\ell) = \bigcup_{u \in Z_1} F_u \quad \text{où} \quad F_u = \{Z_{n-|\ell|-1}(\Theta_u(\cdot), \gamma_{[\ell, \sigma(u)]}) > 0\} \quad (3.11)$$

et d'après la proposition 1

$$P_{\tau(\ell)}^{F_1}(F_u) = P_{i+|\ell|+1, \sigma(u)}(Z_{n-|\ell|-1}(\gamma_{[\ell, \sigma(u)]}) > 0) = p_\ell(\sigma(u)) \quad (3.12)$$

et d'après l'indépendance conditionnelle

$$P_{\tau(\ell)}^{F_1}(B(\ell)^c) = \prod_{u \in Z_1} [1 - p_\ell(\sigma(u))]$$

d'où

$$P_{\tau(\ell)}^N(B(\ell)^c) = (1-\pi)^N \quad \text{avec} \quad \pi = \int_{\mathbb{R}} p_\ell(y) \mu_{\tau(\ell)}^\circ(dy) = \frac{r(\ell)}{m_{\tau(\ell)}} \quad (3.13)$$

En déconditionnant on obtient

$$1-p(\ell) = E_{\tau(\ell)}(1-\pi)^N = g_{\tau(\ell)}(1-\pi) \quad (3.14)$$

1) a) découle alors de l'identité classique $g(s) = 1-m(1-s) + \rho(s)(1-s)$ (l'usage de ρ remplace celui de l'inégalité de Bonferonni [8] nécessitant des moments d'ordre 2, voir aussi plus bas le lemme 4)

Pour 1) b) l'inégalité évidente $k P(\bigcup_{i=1}^k C_i) \geq \sum_{i=1}^k P(C_i)$ jointe à (3.11) donne :

$$N P_{\tau(\ell)}^{F_1}(B(\ell)) \geq 1_{N>0} Z_1(p_\ell)$$

puis

$$P_{\tau(\ell)}^N(B(\ell)) \geq 1_{N>0} \pi$$

et en déconditionnant

$$p(\ell) \geq \beta_{\tau(\ell)} r(\ell)$$

Pour montrer le 2) on remarque d'abord que l'inégalité de Markov donne

$$|\ell| \leq n-1 \Rightarrow p(\ell) \leq q(\ell)$$

d'où $|\ell| \leq n-1 \Rightarrow r(\ell) \leq q(\ell)$

(pour $|\ell| \leq n-2$ on a $r(\ell) \leq \int_{\mathbb{R}} q(y) Q(\bar{\tau}(\ell), dy) = q(\ell)$ par itération de Q et pour $|\ell| = n-1$ on a $r(\ell) = Q(\bar{\tau}(\ell), \gamma_\ell) = q(\ell)$).

D'après a) et b), à ℓ fixé la décroissance de $\rho_{\bar{\tau}(\ell)}^-$ entraîne

$$p(\ell) \geq r(\ell) \left\{ \left[1 - \frac{1}{m_{\bar{\tau}(\ell)}^-} \rho_{\bar{\tau}(\ell)}^- \left(1 - \frac{q(\ell)}{m_{\bar{\tau}(\ell)}^-} \right) \right] \vee \beta_{\bar{\tau}(\ell)}^- \right\} \tag{3.17}$$

Cette relation (évidente hors de \tilde{S}) entraîne pour $0 \leq k \leq n-1$

$$p \geq (\eta_k \vee \beta) r \quad \text{sur } \mathbb{R}^k \tag{3.18}$$

Dans ce cas si on note $a_k = \inf_{\ell \in \tilde{S} \cap \mathbb{R}^k} \frac{p(\ell)}{q(\ell)}$ on a, pour $k \leq n-2$ et $\ell \in \tilde{S} \cap \mathbb{R}^k$

$$p(\ell) \geq (\eta_k \vee \beta) \int_{\mathbb{R}} Q(\bar{\tau}(\ell), dy) p_\ell(y) \geq a_{k+1} (\eta_k \vee \beta) \int_{\mathbb{R}} Q(\bar{\tau}(\ell), dy) q_\ell(y)$$

c'est-à-dire $a_k \geq (\eta_k \vee \beta) a_{k+1}$.

L'égalité $q=r$ sur \mathbb{R}^{n-1} entraîne $a_{n-1} \geq \eta_{n-1} \vee \beta$ ce qui termine la preuve. ▣

Rappelons que notre but est de minorer $P_{o, x_0}(\xi_n(\Delta_u(x)) > 0)$ où par définition

$$\Delta_u(x) = [x-u, x+u]$$

Posons x_0, x, n , étant fixés,

$$\lambda_0 = x_0, \quad \lambda = (\lambda_1, \dots, \lambda_n) \quad \text{avec } \lambda_n = x \quad u, v > 0 \quad \text{et} \tag{3.19}$$

$$\gamma = 1_{\Gamma_\lambda(0, n, u, v)} \quad (\text{voir } \S \text{ II})$$

Dans ces conditions on a

$$P_{o, x_0}(\zeta_n(\Delta_u(x)) > 0) \geq P_{o, x_0}(Z_n(\gamma) > 0) = p(\phi)$$

Pour poursuivre l'étude, vu la proposition précédente il nous faut

1) minorer $q(\phi)$ - ce qui est possible grâce au lemme 2

2) pour tout k minorer η_k et par conséquent

a) majorer $q(\ell)$ pour $|\ell| = k$

b) comparer $\rho_{\tau}^-(\ell)$ à un ρ fixe.

Les 2 lemmes suivants donnent une réponse à a) et b) .

3. Lemmes techniques.

Lemme 3 :

Avec les notations de II, III 1, 2 et en particulier (3.19), sous l'hypothèse

$$\inf_{0 \leq k < n} H(k, \lambda_k, \lambda_{k+1} - \lambda_k) \geq c > 0 \quad (3.20)$$

la condition $\Delta(0, 2u) \leq 1$ entraîne pour tout k tel que $0 \leq k < n$ et tout ℓ de \mathbb{R}^k

$$q(\ell) \leq d_{n-k}^n \quad (3.21)$$

où on a noté pour $0 \leq j \leq n$

$$d_j^n = \exp \{-ja + b\} \quad (3.22)$$

avec $a = c (1 - \Delta(0, 2u))^2$, $b = 2v + 2(M+1) n \Delta(0, 2u)$ \square

Démonstration :

Si $|\ell| = k$ et si λ' est définie par

$$\lambda'_i = \lambda_i \quad 0 < i < k, \quad \lambda'_k = \ell_k, \quad \lambda'_i = \lambda_i - \lambda_k + \lambda_k \quad n \geq i > k$$

alors on a

$$\gamma_\ell \leq \frac{1}{\Gamma_{\lambda'}(k, n, 2u, 2v)}$$

et donc d'après le lemme 2

$$q(\ell) \leq \exp \{2v + (M+1) n \Delta(0, 2u) - I_1(k, n, \lambda') (1 - \Delta(0, 2u))\}$$

et comme $|\lambda'_i - \lambda_i| \leq u$ et $\lambda'_{i+1} - \lambda'_i = \lambda_{i+1} - \lambda_i$ pour tout $k \leq i < n$ on a facilement, dès que $\Delta(0, 2u)$ est inférieur ou égal à 1 :

$$q(\ell) \leq \exp \{2v + 2(M+1)n \Delta(0,2u) - I_1(k,n,\lambda) (1 - \Delta(0,2u))^2\}$$

puis (3.20) permet de conclure. \square

La condition (3.20) se réfère à une "sous-criticité". ([8]) .

Lemme 4 :

Pour tout $(A, \delta) \in (\mathbb{R}_*^+)^2$, il existe $s_0 \in [0, 1[$ et $\bar{\rho}$ fonction décroissante

de $[s_0, 1]$ dans $[0, \bar{\rho}(s_0)]$ vérifiant $\int_{s_0}^1 \frac{\bar{\rho}(s)}{1-s} ds < \infty$, telle que pour toute variable aléatoire N à valeurs dans \mathbb{N} , la condition

$$E N(\text{Log } N)^{1+\delta} \leq A$$

entraîne

$$\rho(s) \leq \bar{\rho}(s) \text{ sur } [s_0, 1]$$

où, comme d'habitude $\rho(s) = E(N - \frac{1-s^N}{1-s})$

Démonstration :

ρ s'écrit aussi $\rho(s) = \sum_{n=0}^{+\infty} P(N > n) (1-s)^n$ ce qui donne pour tout n_0 de \mathbb{N} :

$$\rho(s) \leq n_0 (EN) (1-s) + E(N 1_{N \geq n_0})$$

ou encore

$$A^{-1} \rho(s) \leq n_0 (\text{Log } 2)^{-1-\delta} (1-s) + (\text{Log } n_0)^{-1-\delta}$$

Pour optimiser en n_0 , on pose $\theta = (\text{Log } 2)^{-1-\delta} (1-s)$ et sur $\mathbb{R}^+ \times [1, +\infty)$

$f(\theta, \xi) = \theta \xi + (\text{Log } \xi)^{-1-\delta}$. Le minimum de $f(\theta, \cdot)$ est atteint en un point

$\xi_0(\theta)$. On vérifie alors que le choix $s_0 = (1 - \frac{1+\delta}{2} (\text{Log } 2)^{-2-2\delta})^+$,

$\bar{\rho}(s) = A f(n_0(\theta), \theta)$ où $\theta = (\text{Log } 2)^{-1-\delta} (1-s)$ et $n_0(\theta) = 2v[\xi_0(\theta)]$ convient.

§ IV - PROBABILITES DE PRESENCE - RESULTATS ASYMPTOTIQUES.

On va se donner des hypothèses sur le processus ponctuel de base, construire la famille indexée par le paramètre d'échelle ε puis démontrer les 2 théorèmes limites.

Soit $P_{t,x}$, $(t,x) \in \mathbb{R}^+ \times \mathbb{R}$, une famille mesurable de probabilités sur $M, \mathcal{B}(M)$ telle que pour tout (t,x) , $P_{t,x}$ ait une espérance notée $\mu_{t,x}$. Si ν est l'élément générique de M , on note $N = \nu(1)$.

1. Hypothèses.

a) Hypothèses sur la famille $\mu_{t,x}$, $(t,x) \in \mathbb{R}^+ \times \mathbb{R}$

Elles sont voisines de certaines de celles de Venttsel [12], [4]

$$H0) \quad \forall (t,x,\theta,\nu) \in \mathbb{R}^+ \times \mathbb{R}^3 \quad L(t,x,\theta) = \text{Log} \int e^{\theta \xi} \mu_{t,x} (d\xi) \text{ et} \\ h(t,x,\nu) = \sup_{\lambda} \lambda \nu - L(t,x,\lambda) \text{ sont finies.}$$

$$H1) \quad \text{Il existe } M \geq 0 \text{ tel que } \inf_{t,x,\nu} h(t,x,\nu) \geq -M$$

H2) h est de classe C^1 en t,x , vérifie H0, H1 et il existe C mesurable de \mathbb{R}^+ dans \mathbb{R}^+ , localement bornée, telle que pour tout (t,x,ν) de $\mathbb{R}^+ \times \mathbb{R}^2$ on ait

$$\left| \frac{\partial h}{\partial t} (t,x,\nu) \right| + \left| \frac{\partial h}{\partial x} (t,x,\nu) \right| \leq C(t) (M+1+h(t,x,\nu))$$

H3) Pour tout $T > 0$, il existe h_1 croissante de \mathbb{R}^+ dans \mathbb{R}^+ telle que

$$\forall r > 0, \quad \forall t \leq T, \quad \forall \nu \quad |v| \leq r \Rightarrow \left| \frac{\partial h}{\partial v} (t,x,\nu) \right| \leq h_1(r)$$

H4) Il existe h_2 décroissante de \mathbb{R}^+ dans \mathbb{R}_*^+ telle que

$$\forall r > 0 \quad \forall t,x,\nu \quad |v| \leq r \Rightarrow \frac{\partial^2 h}{\partial v^2} (t,x,\nu) \geq h_2(r)$$

b) Hypothèses sur la loi de N .

H5) Pour tout compact K de $\mathbb{R}^+ \times \mathbb{R}$, il existe c_1, c_2, c_3 et δ dans \mathbb{R}_*^+ tels que

$$(t,x) \in K \quad \Rightarrow \quad c_1 \leq P_{t,x} (N > 0), \quad c_2 \geq E_{t,x} N (\text{Log } N)^{1+\delta}$$

2. Construction de la famille de processus (modèle avec paramètre d'échelle).

Nous allons modéliser maintenant l'idée de reproductions de plus en plus fréquentes avec sauts de plus en plus petits.

Pour tout $(\varepsilon, x) \in \mathbb{R}_*^+ \times \mathbb{R}$ soit $S_{\varepsilon, x}$ l'application de $K(\mathbb{R})$ dans lui-même définie par $S_{\varepsilon, x} f(z) = f(x + \varepsilon z)$. Elle induit sur M une application, notée encore $S_{\varepsilon, x}$ définie par $S_{\varepsilon, x} \nu(f) = \nu(S_{\varepsilon, x} f)$. Pour tout i de \mathbb{N} soit $\bar{P}_{i, x}^\varepsilon$, l'image de $P_{i, x}$ par $S_{\varepsilon, x}$. Pour tout ε , la proposition 1 permet, à partir des $\bar{P}_{i, x}^\varepsilon$ de construire une famille $P_{i, x}^\varepsilon$, $(i, x) \in \mathbb{N} \times \mathbb{R}$ de lois sur (Ω, I) .

3. Probabilité de présence - Minoration asymptotique relative à un chemin.

Le théorème suivant relie la probabilité de présence près de x à l'intégrale d'action de φ , chemin conduisant à x .

Théorème 1 :

Si pour tout ε le processus de branchement spatial construit au 2 est découplé et vérifie H_0, \dots, H_5 alors pour tout (t, x_0, x_1) de $\mathbb{R}^+ \times \mathbb{R}^2$ et φ de $C^1([0, t], \mathbb{R})$ tel que $\varphi(0) = x_0$, $\varphi(t) = x_1$, il existe σ positif tel que l'inégalité

$$(*) \quad \inf_{\substack{k, \varepsilon \\ 0 \leq k\varepsilon < t - \varepsilon}} h(k\varepsilon, \varphi(k\varepsilon), \frac{1}{\varepsilon} (\varphi((k+1)\varepsilon) - \varphi(k\varepsilon))) > 0 \quad (4.1)$$

entraîne

$$\lim_{\substack{\varepsilon \rightarrow 0 \\ u > \sigma\sqrt{\varepsilon}}} \varepsilon \text{Log } P_{0, x_0}^\varepsilon \left[\zeta_{\left[\frac{t}{\varepsilon} \right]} (\Delta_{x_1}^u) > 0 \right] \geq -I(0, t, \varphi) \quad \square$$

Démonstration :

t, x_0, x_1 et φ sont fixés. Comme il a été remarqué avant le lemme 3, si on note $\lambda_s = \varphi(s\varepsilon)$ pour $0 \leq s \leq n = \left[\frac{t}{\varepsilon} \right]$ et $\gamma = 1_{\Gamma_\lambda} (0, n, u, v)$ pour v à choisir plus tard on a :

$$P_{0, x_0}^\varepsilon (\zeta_n (\Delta_u(x)) > 0) \geq P_{0, x_0}^\varepsilon (Z_n(\gamma) > 0) = p(\phi) \quad (4.2)$$

Pour utiliser les lemmes précédents nous avons besoin de quelques évaluations.

a) Evaluations relatives au II :

On vérifie facilement que $\bar{p}_{k,x}^\varepsilon$ a une espérance notée ${}_\varepsilon Q(k,x,\cdot)$ telle que

$${}_\varepsilon Q(k,x,f) = \int_{\mathbb{R}} f(x+\varepsilon z) \mu_{k\varepsilon,x}(dz)$$

ce qui entraîne, pour les quantités définies en II, si $(k,x,\theta,a) \in \mathbb{N} \times \mathbb{R}^3$

$${}_\varepsilon G(k,x,\theta) = L(k\varepsilon,x,\varepsilon\theta) \quad , \quad {}_\varepsilon H(k,x,a) = h(k\varepsilon,x,\frac{a}{\varepsilon}) \quad (4.3)$$

si on note $B = \frac{1}{2} \exp(2C(t))$ l'hypothèse H2) permet d'écrire :

$$\Delta(j,u) \leq B (j\varepsilon+u) \quad \text{pour} \quad (j\varepsilon) \vee u \leq 1 \quad (4.4)$$

et pour tout (k,ξ) de $\mathbb{N} \times \mathbb{R}$, d'après

$$\theta(k,\xi) = \frac{1}{\varepsilon} \frac{\partial h}{\partial v} (k\varepsilon,\xi, \frac{1}{\varepsilon} (\lambda_{k+1} - \lambda_k)) \quad (4.5)$$

et donc grâce à H3) si $\|\dot{\varphi}\|$ désigne la norme uniforme de la dérivée de φ sur $[0,t]$

$$\sup_{\substack{\xi \\ k \leq n-1}} |\theta(k,\xi)| \leq \frac{1}{\varepsilon} h_1 (\|\dot{\varphi}\|) \quad (4.6)$$

Pour le calcul de D_2 et D_2' il faut majorer $\frac{\partial^2}{\partial \theta^2} ({}_\varepsilon G(k,\xi,\theta))$ uniformément en (k,ξ) pour $\theta = \theta(k,\xi)$ ou encore, grâce aux calculs de dualité habituels, majorer $\varepsilon^2 (\frac{\partial^2}{\partial v^2} h(k\varepsilon,\xi, \frac{1}{\varepsilon} (\lambda_{k+1} - \lambda_k)))^{-1}$. H6) fournit le majorant $\varepsilon^2 h_2(\|\dot{\varphi}\|)^{-1}$ et dans ces conditions on a :

$$D_2(0,n,\lambda) \leq D_2 \varepsilon^2 \quad , \quad D_2'(0,n,\lambda) \leq D_2 \quad (4.7)$$

$$\text{où} \quad D_2 = (h_2(\|\dot{\varphi}\|))^{-1} (1 \vee (h_1(\|\dot{\varphi}\|))^2)$$

Pour $k < n$, l'inégalité de Jensen et H2) donnent, si $(1 + \|\dot{\varphi}\|)\varepsilon \leq 1$

$$\begin{aligned} \varepsilon I_1(k,n-k,\lambda) &\leq (1 + B \varepsilon (1 + \|\dot{\varphi}\|)) I(k\varepsilon,t,\varphi) \\ &\quad + B (1 + \|\dot{\varphi}\|) (M+1) t\varepsilon \end{aligned} \quad (4.8)$$

La minoration de $E_{0,x_0}^\varepsilon Z_n(\gamma) = q(\varphi)$ devient, grâce au lemme 2 :

$$q(\phi) \geq (1-tD_2 \left(\frac{\varepsilon}{u^2} + \frac{1}{\varepsilon v^2} \right))^+ \exp - \frac{1}{\varepsilon} D^\varepsilon(0,t,\varphi,u,v) \quad (4.9)$$

$$\begin{aligned} D^\varepsilon(0,t,\varphi,u,v) &= (1+Bu) (1+B\varepsilon (1+ \|\dot{\phi}\|)) I(0,t,\varphi) \\ &\quad + B(M+1) (1+ \|\dot{\phi}\|) (1+Bu) t\varepsilon \\ &\quad + \varepsilon v + B(M+1) t u \end{aligned} \quad (4.10)$$

Il est temps de choisir

$$u = \varepsilon v = 2 \sqrt{tD_2\varepsilon} \quad (4.11)$$

ce qui permet d'obtenir

$$1-tD_2 \left(\frac{\varepsilon}{u^2} + \frac{1}{\varepsilon v^2} \right) = \frac{1}{2} \quad \text{et} \quad \lim_{\varepsilon \rightarrow 0} D^\varepsilon(0,t,\varphi,u,v) = I(0,t,\varphi) \quad (4.12)$$

D'après le lemme fondamental, il nous reste à minorer $\prod_{k=0}^{n-1} (\eta_k v \beta)$.

b) Evaluations relatives au III - Minoration de $\prod_{k=0}^{n-1} (\eta_k v \beta)$.

L'hypothèse (*) correspond à (3.20) et avec les choix précédents, $\Delta(0,2u)$ tend vers 0 avec ε et de plus on a

$$a = c (1-2Bu)^2 = c(1- 4c \sqrt{D_2 t \varepsilon}) \quad (4.13)$$

$$eb = 2\varepsilon v + 4(M+1) tu = 4 \sqrt{D_2 t \varepsilon} (1+2(M+1)t)$$

ce qui permet de conclure que d_n^n tend vers 0 quand ε tend vers 0.

Ici $\bar{\tau}(\phi) = (0, x_0)$ et si on pose $\Lambda = [x_0 - t \|\dot{\phi}\| - 1, x_0 + t \|\dot{\phi}\| + 1]$ on a $S \subset \Lambda^n$ et donc $\bar{\tau}(\ell) \in [0, \dots, n] \times \Lambda$ si $\ell \in \tilde{S}$. Si K désigne le compact $[0, t] \times \Lambda$ on a donc, d'après H5) et H1)

$$\sup_{\ell \in \tilde{S}} \beta_{\bar{\tau}(\ell)}^- \geq e^{-M} c_1 \stackrel{\text{déf}}{=} \beta \quad \text{et} \quad \sup_{\ell \in \tilde{S}} E_{\bar{\tau}(\ell)}^- N(\text{Log } N)^{1+\delta} \leq c_2$$

Le lemme nous fournit un couple $s_0, \bar{\rho}$ tel que

$$\int_{s_0}^1 \frac{\bar{\rho}(s)}{1-s} ds < \infty \quad \text{et} \quad \sup_{\ell \in \tilde{S}} \rho_{\bar{\tau}(\ell)}^-(s) \leq \bar{\rho}(s) \quad \text{pour } s \in [s_0, 1]$$

On peut toujours augmenter s_0 pour que $\bar{\rho}(s_0)$ soit inférieur à $\frac{c_1}{2}$. D'après ce qui a été vu plus haut, il existe ε_0 tel que

$$\varepsilon \leq \varepsilon_0 \Rightarrow d_n^n \leq c_1 (1-s_0)$$

On pose alors

$$j_n = \inf \{j : 1 \leq j \leq n \text{ et } d_j^n \leq c_1 (1-s_0)\} \quad (4.14)$$

ce qui entraîne pour $0 \leq k \leq n-j_n$

$$\eta_k \geq 1 - \frac{1}{c_1} \bar{\rho} \left(1 - \frac{1}{c_1} d_{n-k}^n\right)$$

et par suite

$$\prod_{k=0}^{n-1} (\eta_k \vee \beta) \geq \beta^{j_n-1} \prod_{j=j_n}^n \left(1 - \frac{1}{c_1} \bar{\rho} \left(1 - \frac{d_j^n}{c_1}\right)\right) \quad (4.15)$$

L'inégalité $\text{Log}(1-x) \geq -2x \text{Log } 2$ valable pour $0 \leq x \leq 1/2$ donne alors

$$\prod_{k=0}^{n-1} (\eta_k \vee \beta) \geq \beta^{j_n-1} \exp - \left\{ \frac{2 \text{Log } 2}{c_1} \sum_{j=j_n}^n \bar{\rho} \left(1 - \frac{d_j^n}{c_1}\right) \right\} \quad (4.16)$$

La définition des d_j^n et la décroissance de $\bar{\rho}$ permet de majorer la dernière somme par $\frac{c_1}{2} + c_1 \int_{s_0}^1 \frac{\bar{\rho}(s)}{1-s} ds$, constante ne dépendant que de c_1 et c_2 .

On peut désormais conclure en justaposant (4.2), (3.10), (4.9), (4.16) et en

remarquant que j_n-1 , qui est majoré par $\frac{1}{a} \{b - \text{Log } c_1 (1-s_0)\}$ vérifie

$\lim_{\varepsilon \rightarrow 0} \varepsilon(j_n-1) = 0$, par la formule

$$\lim_{\varepsilon \rightarrow 0} \frac{\varepsilon \text{Log } P_{0,x_0}^\varepsilon \left(\zeta_{\left[\frac{t}{\varepsilon}\right]} (\Delta_u(x)) > 0 \right)}{u=2\sqrt{tD_2\varepsilon}} \geq -I(0,t,\varphi) \quad \square$$

Remarque :

On ne peut pas en général remplacer δ par 0 dans H5) mais si le processus est homogène ($\rho_{t,x} \equiv \rho$ pour tout t,x) on a $\int_0^1 \frac{\rho(s)ds}{1-s} \leq E N(1 + \text{Log } N)$ et la conclusion reste vraie en remplaçant δ par 0.

4. Probabilité de présence - Minoration définitive.

On voudrait naturellement optimiser en φ le théorème précédent, mais $\sigma = 2\sqrt{tD_2}$ dépend de $\|\dot{\varphi}\|$. Nous allons montrer qu'avec deux hypothèses supplémentaires sur h , $\inf I(0,t,\varphi)$ est atteint. L'une H6) (dite condition de croissance) vise, par la compacité des ensembles de niveaux de I à montrer que l'inf est atteint dans la classe des chemins absolument continus ($A[0,t]$), l'autre H7) assure la régularité (au moins C^1) des trajectoires optimales.

H6) Il existe \underline{h} de \mathbb{R}^+ dans \mathbb{R} vérifiant

$$\lim_{r \rightarrow +\infty} r^{-1} \underline{h}(r) = +\infty \quad \text{et} \quad \forall (t,x,v) \in \mathbb{R}^+ \times \mathbb{R}^2 \quad \underline{h}(|v|) \leq h(t,x,v)$$

Elle est équivalente à

H6') $\forall \theta \in \mathbb{R} \quad \sup_{t,x} L(t,x,\theta) < \infty$

On montre facilement que \underline{h} peut être choisie convexe croissante. H6) entraîne H1) donc l'existence de M .

H7) Il existe C , comme dans H2) vérifiant

$$\forall (t,x,v) \in \mathbb{R}^+ \times \mathbb{R}^2 \quad \left| \frac{\partial h}{\partial v}(t,x,v) \right| \leq C(t) [M+1+h(t,x,v)]$$

On rappelle que A est l'ensemble des fonctions absolument continues et I l'intégrale d'action.

Proposition 2 :

Sous H0), H2), H6), H7) pour tous $0 < t_0 < t_1$ et x_0, x_1 de \mathbb{R} , $\inf I(t_1, t_2, \varphi)$ pris sur les φ de $A[t_0, t_1]$ telles que $\varphi(t_0) = x_0$, $\varphi(t_1) = x_1$ est atteint pour une φ de classe C^1 . \square

Ce problème, bien connu, est traité dans de nombreux ouvrages, mais comme les différentes preuves font appel à des hypothèses différentes qui se recoupent sans se recouvrir, on rappellera ici comment nos hypothèses, qui sont loin d'être

minimales, permettent d'aboutir (H6)) à l'existence et (H2) + H7)) à la régularité de chemins optimaux.

Démonstration :

On omettra parfois t_0 et t_1 (fixés) dans $I(t_0, t_1, \cdot)$. On peut toujours supposer $h \geq 0$ quitte à remplacer h par $h+M$. Comme I est positive sur A , l'inf existe. Si φ_n est minimisante, il existe $c > 0$ et n_0 tels que $I(\varphi_n) \leq c$ pour $n \geq n_0$. D'après H6) ceci entraîne l'équi-intégrabilité de la suite $\dot{\varphi}_n$ et donc sa faible relative compacité dans L^1 . On extrait une sous-suite, notée encore $\dot{\varphi}_n$ convergeant faiblement vers u . La fonction φ , définie sur $[t_0, t_1]$ par $\varphi(t) = x_0 + \int_0^t u(s) ds$ est dans $A(\dot{\varphi} = u)$ et est limite simple de φ_n . D'après un résultat classique ([14], p. 243) pour tout (t, x) la convexité de $h(t, x, \cdot)$, jointe à H6) entraîne

$$\int_{t_0}^{t_1} h(t, \varphi(t), u(t)) dt \leq \underline{\lim} I(\varphi_n)$$

et donc φ réalise l'inf.

La régularité se prouve avec les équations d'Euler. Si $B \subset A$ est l'ensemble des ψ telles que $\dot{\psi} \in L_\infty$ et $\psi(t_0) = \psi(t_1) = 0$, alors pour $(\varphi, \psi, \lambda) \in A \times B \times \mathbb{R}$ on peut écrire

$$\lambda^{-1} \{I(\varphi + \lambda\psi) - I(\varphi)\} = \int_{t_0}^{t_1} f(\lambda, t) dt$$

$$\text{où } f(\lambda, t) = \lambda^{-1} \{h(t, \varphi(t) + \lambda\psi(t), \dot{\varphi}(t) + \lambda\dot{\psi}(t)) - h(t, \varphi(t), \dot{\varphi}(t))\}$$

vérifie

$$\lim_{\lambda \rightarrow 0} f(\lambda, t) = \psi(t) \frac{\partial h}{\partial x}(t, \varphi(t), \dot{\varphi}(t)) + \dot{\psi}(t) \frac{\partial h}{\partial v}(t, \varphi(t), \dot{\varphi}(t))$$

et d'après H2) + H7) il existe $B > 0$ et $r > 0$ tels que sur $[t_0, t_1]$

$$|\lambda| \leq r \Rightarrow f(\lambda, t) \leq B (|\psi(t)| + |\dot{\psi}(t)|) (1 + h(t, \varphi(t), \dot{\varphi}(t)))$$

Ce majorant étant intégrable (ψ et $\dot{\psi}$ sont bornés), le théorème de Lebesgue

permet de dire que si $\varphi \in A$ réalise l'inf alors pour tout ψ de B on a :

$$\int_{t_0}^{t_1} \{ \psi(t) \frac{\partial h}{\partial x} (t, \varphi(t), \dot{\varphi}(t)) + \dot{\psi}(t) \frac{\partial h}{\partial v} (t, \varphi(t), \dot{\varphi}(t)) \} dt = 0$$

ou encore, en intégrant par parties

$$\int_{t_0}^{t_1} \dot{\psi}(t) \left\{ \frac{\partial h}{\partial v} (t, \varphi(t), \dot{\varphi}(t)) - \int_{t_0}^t \frac{\partial h}{\partial x} (\sigma, \varphi(\sigma), \dot{\varphi}(\sigma)) d\sigma \right\} dt = 0$$

On en déduit qu'il existe a réel tel que (Euler sous forme intégrale)

$$\frac{\partial h}{\partial v} (t, \varphi(t), \dot{\varphi}(t)) - \int_{t_0}^t \frac{\partial h}{\partial x} (\sigma, \varphi(\sigma), \dot{\varphi}(\sigma)) d\sigma = a \quad \text{p.p. sur } [t_0, t_1]$$

si on pose $p(t) = a + \int_{t_0}^t \frac{\partial h}{\partial x} (\sigma, \varphi(\sigma), \dot{\varphi}(\sigma)) d\sigma$ on a p.p.

$$p(t) = \frac{\partial h}{\partial v} (t, \varphi(t), \dot{\varphi}(t)) .$$

Grâce à la dualité, la formule précédente s'inverse en :

$$\dot{\varphi}(t) = \frac{\partial L}{\partial \theta} (t, \varphi(t), p(t)) \quad \text{p.p.}$$

(si les h n'étaient pas des transformées de Cramer, il faudrait supposer la stricte convexité de h en v).

φ est continue et p aussi par construction donc $\dot{\varphi}$ coïncide p.p. avec une fonction continue. Donc φ est de classe C^1 (et de plus vérifie les équations d'Euler :

$$\dot{p}(t) = \frac{\partial h}{\partial x} (t, \varphi(t), \dot{\varphi}(t)) \quad p(t) = \frac{\partial h}{\partial v} (t, \varphi(t), \dot{\varphi}(t)) \quad) . \quad \square$$

Tout ce qui précède nous a permis en fait de démontrer le théorème suivant. On note $\bar{I}(t_0, t_1, x_0, x_1) = \inf I(t_0, t_1, \varphi)$ pris sur l'ensemble des $\varphi \in A[t_0, t_1]$ tels que $\varphi(t_0) = x_0$, $\varphi(t_1) = x_1$.

Théorème 2 :

Si le processus de branchement spatial construit en IV.2 est découplé et vérifie H0) ... H7) alors pour tout (t, x_0, x_1) de $\mathbb{R}^+ \times \mathbb{R}^2$ tel qu'un des chemins

de classe C^1 atteignant \bar{I} vérifie (*), alors il existe $\sigma > 0$ tel que :

$$\lim_{\substack{\varepsilon \rightarrow 0 \\ u > \sigma\sqrt{\varepsilon}}} \varepsilon \text{Log } P_{0,x_0}^\varepsilon \left(\zeta_{\left[\frac{t}{\varepsilon}\right]}(\Delta_{x_1}^u) > 0 \right) \geq -\bar{I}(0,t,x_0,x_1) \quad \square$$

BIBLIOGRAPHIE :

- [1] AZENCOTT (R.) RUGET (G.), *Mélanges d'équations différentielles et grands écarts à la loi des grands nombres. Zeits. für Wahr.*, 1977, t. 38, p. 1-54.
- [2] BIGGINS (J.D.), *Chernoff's theorem in the branching random walk, Journ. of Appl. Prob.*, 1977, t. 14, p. 630-636.
- [3] BIGGINS (J.D.), *Growth rates in the branching random walk, Zeits. für Wahr.*, 1979, t. 48, p. 17-34.
- [4] FREIDLIN (M.I.), VENTSEL (A.D.), *Random perturbations of dynamical systems, Springer, Berlin 1984.*
- [5] JAGERS (P.), *Branching processes with biological applications. Wiley, London 1975.*
- [6] LAREDO (C.), ROUAULT (A.), *Grandes déviations, dynamique de populations et phénomènes malthusiens, Ann. de l'IHP*, 1983, vol. XIX, n° 4, p.323-350.
- [7] NEVEU (J.), *Arbres et processus de Galton-Watson (à paraître).*
- [8] ROUAULT (A.), *Comportement asymptotique d'un processus de branchement spatial en zone sous-critique, Comptes-Rendus de l'Ac. d. Sc.* 1984, n° 299, série I, p. 887-890.
- [9] ROUAULT (A.), *Probabilités de présence dans un processus de branchement spatial markovien (soumis aux C.R.A.S.).*
- [10] ROUAULT (A.), *Espérances et majorations pour un processus de branchement spatial markovien (soumis aux Ann. IHP), voir Chap. G.*
- [11] RUGET (G.), *Large deviations and more or less rare events in population dynamics. Colloque Luminy. Lecture Notes in Biom. n° 49 (1981), p. 388-400 .*
- [12] VENTSEL (A.D.), *Rough limit theorems on large deviations for Markov stochastic processes. Theor. of Prob. and its Appl.* 1976, t. 21, n° 2, p. 227-242 et t. 21 n° 3, p. 499-512.

- [13] *AZENCOTT (R.), Cours de l'Ecole d'Eté de Saint-Flour, Lecture Notes in Math. n° 774 (1978).*

- [14] *EKELAND (I.), TEMAM (R.), Convex analysis and variationnal problems North-Holland, Amsterdam, 1976.*

- [15] *MATTHES (K.), KERSTAN (J.), MECKE (J.), Infinitely divisible point processes, Wiley, Chichester, 1978.*

G. Espérances et majorations pour un processus de branchement
spatial markovien.

ESPERANCES ET MAJORATIONS POUR UN PROCESSUS DE BRANCHEMENT
SPATIAL MARKOVIEN

Alain ROUAULT

Unité Associée 743-A.D.4
Mathématique (Bât.425)
91405 - ORSAY Cedex

Résumé : - Pour un processus de Galton-Watson $(Z_n)_{n \in \mathbb{N}}$ supercritique, $\frac{1}{n} \text{Log } EZ_n$ et $\frac{1}{n} \text{Log } Z_n$ ont même limite $\text{Log } m$. Ce résultat est étendu dans [2] au cas spatial homogène. Dans le cas markovien [6] avec population initiale exponentielle, on donne ici une limite pour l'analogie de la première expression et une majoration en probabilité pour une analogie de la seconde. Deux formules variationnelles sont en jeu ; la seconde, avec contrainte sur l'intégrale d'action, caractérise par là-même les zones de "présence".

Abstract : - For a supercritical Galton-Watson process $(Z_n)_{n \in \mathbb{N}}$, $\frac{1}{n} \text{Log } EZ_n$ and $\frac{1}{n} \text{Log } Z_n$ have the same limit $\text{Log } m$. This result is extended in [2] to homogeneous spatial case. In the markovian case [6] with exponential initial population, we give here a limit for the analogue of the first expression and an overestimation - in probability - for an analogue of the second one. Two variational formulas are at stake ; the second one, with constraint over action integral, allows the definition of "presence areas".

ESPERANCES ET MAJORATIONS POUR UN PROCESSUS DE BRANCHEMENT
 SPATIAL MARKOVIEN

Alain ROUAULT

INTRODUCTION GENERALE :

On continue ici l'étude des processus de branchement spatiaux markoviens commencée en [6], dont les principales conjectures ont été présentées en [7].

On se donne une famille indexée par $\mathbb{R}^+ \times \mathbb{R}$ (temps, espace) de lois de probabilité sur M , espace des mesures ponctuelles finies sur \mathbb{R} . L'élément générique est $P_{t,x}(dv)$, ayant une espérance notée $\mu_{t,x}$, mesure de Radon positive sur \mathbb{R} . $P_{t,x}$ décrit, à l'échelle 1, la loi des positions des enfants d'un individu situé en x à l'instant t . Le cas homogène correspond à $P_{t,x} \equiv P$ pour tout t, x .

On construit une famille indexée par ε tendant vers 0 de processus de branchement spatiaux. Pour ce faire, on se donne $x_0 \in \mathbb{R}$, position d'un ancêtre (individu de la génération 0). Il vit un temps ε puis est remplacé par des enfants dont les positions forment un élément (aléatoire) de M obtenu de la manière suivante : on tire un élément de M suivant P_{ε, x_0} , on lui applique une homothétie de rapport ε puis une translation de vecteur x_0 . Les enfants de l'ancêtre forment la première génération et sont datés de l'instant ε de sorte qu'un tel individu, situé en x , sera à l'instant 2ε remplacé par des enfants dont les positions formeront un élément de M obtenu par tirage suivant $P_{2\varepsilon, x}$, ε -homothétie et x -translation. On itère la procédure, mais (nouveau par rapport à [6]) à partir d'une population initiale de taille $\exp \frac{1}{\varepsilon} f_0(x)$ au voisinage de

x , pour tout x d'un compact F_0 où $f_0 \geq 0$. On s'intéresse à $\hat{\zeta}_t^\varepsilon$, élément de M représentant les individus en vie à t . L'asymptotique est celle des grandes déviations : reproductions de plus en plus fréquentes (aux instants de $\varepsilon\mathbf{N}$) et les sauts de plus en plus petits (rapport ε).

Dans le cas homogène (et à ancêtre unique) la méthode ([2]) a consisté à évaluer d'abord des espérances puis à comparer les v.a. à leurs espérances. L'espérance de la n -ième génération, égale à μ^{*n} relevait d'un calcul type Chernov. Dans le cas non-homogène on se sert des grandes déviations pour chaînes de Markov ([1], [8]). Si M_1 majore uniformément les taux de croissance, les $M_1^{-1} \mu_{t,x}$ règlent le déplacement markovien d'une seule particule (avec disparition). L'outil fondamental est l'intégrale d'action définie sur $A[t_0, t_1]$, ensemble des fonctions absolument continues de $[t_1, t_2]$ dans \mathbb{R} par

$$I(t_1, t_2, \varphi) = \int_{t_1}^{t_2} h(s, \varphi(s), \dot{\varphi}(s)) ds$$

où $h(t, x, \cdot)$ est la transformée de Cramer de $\mu_{t,x}$. Un chemin φ doit être vu comme la limite (en ε) d'une ligne généalogique constituée par les positions des ancêtres successifs d'un individu.

On montre au III que $\varepsilon \text{Log } E \hat{\zeta}_t^\varepsilon ([x-r(\varepsilon), x+r(\varepsilon)])$ pour un certain $r(\varepsilon)$ tendant vers 0 est voisine de

$$T f_0(0, t, x) = \sup_{\substack{f_0(\varphi(0)) \geq 0 \\ \varphi(t) = x}} f_0(\varphi(0)) - I(0, t, \varphi)$$

La contribution essentielle, stabilisée par l'opération εLog , à l'espérance de la population en x provient de lignes généalogiques dans un tube autour d'un chemin φ (dit optimal) réalisant ce sup. S'il existe $s < t$ tel que $f_0(\varphi(0)) - I(0, s, \varphi) < 0$, alors par référence au modèle homogène, cela signifie qu'avec une probabilité voisine de 1, la population est éteinte au voisinage de $\varphi(s)$ à l'instant s . La contribution correspondant à φ en x est nulle au temps

t avec la même probabilité donc si on attend $\varepsilon \text{Log} \hat{\xi}_t^\varepsilon ([x-r(\varepsilon), x+r(\varepsilon)] \simeq \alpha(t,x) > 0$, α correspondant à une formule variationnelle ne sera sûrement pas atteint pour ce chemin. En fait la conjecture (cf. [7]) est

$$\alpha(t,x) = \sup_{\substack{f_0(\varphi(0)) \geq 0 \\ \varphi(t) = x \\ \forall s \leq t \quad T f_0(0,s,\varphi(s)) \geq 0}} f_0(\varphi(0)) - I(0,t,\varphi)$$

Il va être nécessaire de découper $[0,t]$ en tranches d'épaisseur τ , suffisamment grande pour que le processus ait eu le temps de donner des résultats asymptotiques, mais suffisamment faible pour que qualitativement tout se passe comme pour un processus homogène ($\varepsilon \text{Log} \xi \simeq \varepsilon \text{Log} E\xi$ là où celle-ci est positive). Il est bien entendu que les points de départ spatiaux d'une tranche sont les points d'arrivée de la tranche précédente ayant une densité de population suffisante. L'étude d'une tranche nécessite notamment un encadrement de l'espérance, donc les mêmes outils que dans l'espérance globale. Mais l'empilement d'un nombre croissant ($\simeq \frac{t}{\tau}$ à t fixé) de ces tranches exige de ne pas s'en tenir aux parties principales quand ε tend vers 0, c'est-à-dire d'évaluer les restes de façon uniforme.

Ceci conduit dans I à l'étude des itérées de T en vue d'obtenir des pentes uniformément bornées pour les chemins optimaux. Pour notre modèle précisé en II, nous énonçons et démontrons en III les résultats asymptotiques sur l'espérance du processus. Ceci a notamment pour conséquence, en IV une caractérisation τ -approchée de la zone "désertique" à l'instant t , et une majoration, en probabilité, de $\varepsilon \text{Log} \hat{\xi}_t^\varepsilon$ par le résultat de $[\frac{t}{\tau}]$ itérations de T appliquées à f_0 . Dans un prochain article nous étudierons la minoration correspondante et la convergence de l'algorithme vers α quand $\tau \rightarrow 0$.

§ I - ETUDE D'UN PROBLEME VARIATIONNEL LIE A UN MODELE DE CROISSANCE DE POPULATIONS SPATIALES.

1. Introduction :

Dans le modèle probabiliste étudié, l'évolution entre t_0 et t_1 du logarithme de la taille locale d'une population fait intervenir une fonction f réglant la taille initiale et une fonctionnelle $I(t_0, t_1, \varphi) = \int_{t_0}^{t_1} h(t, \varphi(t), \dot{\varphi}(t)) dt$ définie sur les chemins absolument continus de $[t_0, t_1]$ dans \mathbb{R} (représentant les lignes généalogiques). h est astreinte à une condition de croissance sur-linéaire, de convexité par rapport à la troisième variable et de régularité par rapport aux trois. On est amené à étudier le problème variationnel

$$(*) \quad \sup_{\substack{\varphi: \varphi(t_1) = x \\ f(\varphi(t_0)) \geq 0}} f(\varphi(t_0)) - I(t_0, t_1, \varphi)$$

(f représente le logarithme de la taille d'une population). En posant $g = f$ si $f \geq 0$ et $g = -\infty$ si $f < 0$ on ramène (*) au problème classique de Bolza

$$(**) \quad \sup_{\varphi: \varphi(t_1) = x} g(\varphi(t_0)) - I(t_0, t_1, \varphi)$$

Après avoir énoncé les hypothèses sur h , nous rappellerons au paragraphe 2 les propriétés d'existence et de régularité des chemins optimaux pour le problème de Bolza et nous démontrerons un lemme technique.

Dans un modèle approché, τ est un pas de temps destiné à tendre vers 0 et la population au temps $t = n\tau$ sera décrite par la fonction obtenue en itérant n fois la procédure précédente. Si $Tf(t_0, t_1, x)$ désigne la valeur du sup dans (*) on considère

$$\begin{aligned} & f_0 = f \\ & \dots\dots \\ k \geq 1 & \quad f_k(x) = \begin{cases} Tf_{k-1}((k-1)\tau, k\tau, x) & \text{si } \{f_{k-1} \geq 0\} \neq \emptyset \\ = -\infty & \text{sinon} \end{cases} \end{aligned}$$

Si φ_k désigne un chemin optimal à l'étape k , notre but est d'obtenir une majoration de $\|\dot{\varphi}_k\|_\infty$ uniforme en k, τ vérifiant $k\tau \leq t_1$ fixé. Une propriété de Lipschitz sur f se transmet aux f_k (mais une régularité plus forte se perd dès la 1ère étape). Comme dans la théorie classique (sans contrainte), il y a relation étroite entre les majorants de $\dot{\varphi}_k$, de p_k (lié à φ_k par une équation d'Euler) et le coefficient de Lipschitz de f_{k-1} .

Dans le paragraphe 3, après avoir rappelé le formalisme des sous-différentiels, bien adapté ici, nous donnerons pour le problème à un pas une majoration (uniforme en x) de $|\dot{\varphi}|$ et de $|\partial Tf|$. L'idée fondamentale est la suivante : le problème sans contrainte donne déjà, grâce à la condition de croissance, des chemins optimaux à pentes bornées ; la contrainte introduit une majoration sur les pentes pour d'autres raisons : la saturation a tendance à envoyer $\dot{\varphi}$ sur des régions où h est négatif et $\{u : h(t,x,u) \leq 0\}$ est un compact sous nos hypothèses.

Une itération mécanique des majorations précédentes conduit à $+\infty$. Néanmoins, la dérivabilité p.p. de Tf , conséquence du paragraphe 3 permet, dans le paragraphe 4, un traitement local qui donne par récurrence les majorations désirées pour le problème à n pas. Enfin en 5. on donne un résultat annexe utile pour la suite.

2. Hypothèses - Existence - Régularité :

On dit qu'une fonction F de \mathbb{R} vérifie la condition de croissance si on a

$$(CC) \quad \frac{F(v)}{|v|} \xrightarrow{|v| \rightarrow +\infty} +\infty$$

On considère une fonction h mesurable de $\mathbb{R}^+ \times \mathbb{R}^2$ dans \mathbb{R} , telle que pour tout (t,x) de $\mathbb{R}^+ \times \mathbb{R}$, $h(t,x,\cdot)$ soit convexe et vérifie (CC). On note $L(t,x,\cdot)$ la duale de Young de $h(t,x,\cdot)$. On définit ainsi une fonction L de $\mathbb{R}^+ \times \mathbb{R}^2$ dans \mathbb{R} . Nous utiliserons dans cette partie les hypothèses suivantes. Elles ne sont pas systématiquement minimales ; il semble en particulier que dans

[H3] la lipschitzianité de h suffirait. Mais nous les énonçons ainsi pour des raisons de commodité (cf. [6]).

[H1] $\inf_{t,x} h(t,x,\cdot)$ vérifie (CC), c'est-à-dire qu'il existe \underline{h} de \mathbb{R}^+ dans \mathbb{R} , vérifiant $\frac{\underline{h}(r)}{r} \xrightarrow{r \rightarrow +\infty} +\infty$ (CC') et

$$\inf_{t,x} h(t,x,v) \geq \underline{h}(|v|) \quad \text{pour tout } v.$$

Elle est équivalente à :

[H'1] Pour tout $\theta \in \mathbb{R}$ $\sup_{t,x} L(t,x,\theta) < \infty$.

Dans ce cas on pose

$$M = \sup_{t,x} [L(t,x,0)]^+ , \quad \bar{L}(\theta) = \sup_{t,x} L(t,x,\theta) \vee L(t,x,-\theta) \vee M$$

\bar{L} est convexe, paire, sa duale $(\bar{L})^*$ est convexe, s.c.i., paire, vérifie (CC) donc atteint son minimum M_0 en $\pm r_0$ ($r_0 > 0$). On pose

$$\underline{h}(r) = M_0 \quad \text{pour } 0 \leq r \leq r_0 , \quad \underline{h}(r) = (\bar{L})^*(r) \quad \text{pour } r \geq r_0 .$$

Alors \underline{h} est croissante, convexe, s.c.i., vérifie $-M \leq \bar{L}^*(+r) \leq \underline{h}(r)$ ($r \geq 0$) et convient pour [H1]. On note $\underline{h} = \underline{h} + M > 0$.

[H2] Pour tout $v \in \mathbb{R}$, $\sup_{t,x} h(t,x,v) < +\infty$. On définit alors \bar{h} , convexe par

$$\bar{h}(v) = \sup_{t,x} h(t,x,v)$$

et $\bar{\bar{h}}$, croissante de \mathbb{R}^+ dans \mathbb{R} par

$$\bar{\bar{h}}(r) = \sup_{|v| \leq r} \bar{h}(v)$$

[H3] h vérifie H1 et est de classe C^1 . Pour tout t,x , $h(t,x,\cdot)$ est strictement convexe. Il existe C mesurable de \mathbb{R}^+ dans \mathbb{R}^+ , localement bornée, qu'on peut toujours supposer croissante, telle que pour tout

$$(t,x,v) \in \mathbb{R}^+ \times \mathbb{R}^2$$

$$\left| \frac{\partial h}{\partial t}(t, x, v) \right| + \left| \frac{\partial h}{\partial x}(t, x, v) \right| + \left| \frac{\partial h}{\partial v}(t, x, v) \right| \leq C(t) [M+1+h(t, x, v)]$$

[H2] + [H3] entraîne

[H3'] Pour tout $T > 0$ il existe une fonction croissante h_1 de \mathbb{R}^+ dans \mathbb{R}^+ telle que

$$t \leq T, \quad |v| \leq r \text{ entraîne pour tout } x \quad \left| \frac{\partial h}{\partial v}(t, x, v) \right| \leq h_1(r)$$

[H4] h est de classe C^1 par rapport à v et il existe h_1 strictement croissante de \mathbb{R} dans \mathbb{R} , tendant vers $+\infty$ avec son argument telle que pour tout $(t, x, v) \in \mathbb{R}^+ \times \mathbb{R}^2$

$$h_1(v) \leq \frac{\partial h}{\partial v}(t, x, v)$$

[H3'] + [H4] entraîne

[H4'] Pour tout $T > 0$, il existe L_1 croissante de \mathbb{R}^+ dans \mathbb{R}^+ telle que

$$\forall t \leq T, \quad |\theta| \leq r, \quad x \in \mathbb{R} \quad \left| \frac{\partial L}{\partial \theta}(t, x, \theta) \right| \leq L_1(r)$$

Le résultat principal de ce paragraphe se trouve dans la proposition suivante.

Pour tous $0 < t_0 < t_1 < +\infty$ on note $A[t_0, t_1]$ l'ensemble des fonctions absolument continues de $[t_0, t_1]$ dans \mathbb{R} et pour $\varphi \in A[t_0, t_1]$

$$I(t_0, t_1, \varphi) = \int_{t_0}^{t_1} h(t, \varphi(t), \dot{\varphi}(t)) dt$$

Proposition 1 :

Sous les hypothèses [H1] et [H3] si f est une fonction de \mathbb{R} dans $\mathbb{R} \cup \{-\infty\}$ s.c.s. et vérifiant $-\infty < \sup f < +\infty$, alors

$$\sup_{\substack{\varphi \in A[t_0, t_1] \\ \varphi(t_1) = x}} f(\varphi(t_0)) - I(t_0, t_1, \varphi)$$

est atteint et les chemins optimaux sont de classe C^1 .

Ce problème, bien connu, est traité dans de nombreux ouvrages, mais comme les différentes démonstrations se font sous des hypothèses différentes qui se recoupent sans se recouvrir, nous renvoyons le lecteur à [6] prop. 2 (qu'il faut très légèrement modifier à cause de f). En fait [H1] entraîne l'atteinte dans A et [H3] la régularité des chemins optimaux.

Si $f(x_0)=0$ et $f = -\infty$ sur $\{x_0\}^C$ on retrouve le problème à points fixes (t_0, x_0) , (t_1, x) . Pour f s.c.s. quelconque un φ solution du problème de Bolza est aussi solution du problème à points fixes $(t_0, \varphi(t_0))$, (t_1, x) .

On note désormais pour tout (t_0, t_1, x_0, x_1) de $\mathbb{R}_+ \times \mathbb{R}^2$, $\Phi(t_0, t_1, x_0, x_1)$ l'ensemble des φ de $C^1([t_0, t_1])$ vérifiant

$$\varphi(t_0)=x_0 \quad , \quad \varphi(t_1)=x_1 \quad , \quad I(t_0, t_1, \varphi) = \inf_{\substack{\psi \in A[t_0, t_1] \\ \psi(t_0)=x_0 \\ \psi(t_1)=x_1}} I(t_0, t_1, \psi)$$

c'est-à-dire l'ensemble des chemins optimaux pour le problème à points fixes.

Nous énonçons et prouvons maintenant un lemme technique.

Lemme 1 :

Définissons, pour $(a, b, t^1) \in \mathbb{R} \times \mathbb{R}_+^2$ les sous-ensembles suivants de $\mathbb{R}_+^2 \times \mathbb{R}^2$

$$\Delta_1(a) = \left\{ (t, \tau, x_0, x_1) : \Phi(t, t+\tau, x_0, x_1) \subset \{ \varphi : I(t, t+\tau, \varphi) \leq a \} \right\}$$

$$\Delta_2(a) = \left\{ (t, \tau, x_0, x_1) : \Phi(t, t+\tau, x_0, x_1) \subset \{ \varphi : I(t, t+\tau, \varphi) \leq a\tau \} \right\}$$

$$\Delta_3(b) = \left\{ (t, \tau, x_0, x_1) : |x_1 - x_0| \leq b\tau \right\}$$

$$\Delta_4(b) = \left\{ (t, \tau, x_0, x_1) : \varphi \in \Phi(t, t+\tau, x_0, x_1), \quad s \in [t, t+\tau] \Rightarrow |\dot{\varphi}(s)| \leq b \right\}$$

$$\Delta_5(t^1) = \left\{ (t, \tau, x_0, x_1) : t+\tau \leq t^1 \right\}$$

Sous $H1, H2, H3, H4$ on a :

$$1) \quad \forall a \in \mathbb{R} \quad , \quad \exists b \in \mathbb{R}_+ \quad : \quad \Delta_2(a) \subset \Delta_3(b) \quad (1.1)$$

$$2) \quad \forall (a, b, t^1) \in \mathbb{R} \times \mathbb{R}_+^2 \quad \exists c \in \mathbb{R}_+ \quad : \quad \Delta_1(a) \cap \Delta_3(b) \cap \Delta_5(t^1) \subset \Delta_4(c) \quad (1.2)$$

Preuve du lemme 1 :

1) D'après [H1] , si $(t, \tau, x_0, x_1) \in \Delta_2(a)$ et $\varphi \in \Phi(t, \tau, x_0, x_1)$ on a

$$\tau^{-1} \int_t^{t+\tau} \underline{h} (|\dot{\varphi}(s)|) ds \leq a \quad (1.3)$$

La croissance de \underline{h} , jointe à sa convexité (Jensen) donne alors :

$$\underline{h} \left(\frac{|x_1 - x_0|}{\tau} \right) = \underline{h} \left(\tau^{-1} \left| \int_t^{t+\tau} \dot{\varphi}(s) ds \right| \right) \leq \underline{h} \left(\tau^{-1} \int_t^{t+\tau} |\dot{\varphi}(s)| ds \right) \leq \tau^{-1} \int_t^{t+\tau} \underline{h} (|\dot{\varphi}(s)|) ds$$

d'où $\underline{h} \left(\frac{|x_1 - x_0|}{\tau} \right) \leq a$ et on conclut à l'existence de b car \underline{h} vérifie CC .

2) Soit $(t, \tau, x_0, x_1) \in \Delta_3(b)$. Tout φ de $\Phi(t, \tau, x_0, x_1)$ vérifie les équations d'Euler partout : il existe p de classe C^1 tel que pour tout s de $[t, t+\tau]$ on ait :

$$\dot{p}(s) = \frac{\partial h}{\partial x} (s, \varphi(s), \dot{\varphi}(s)) \quad , \quad p(s) = \frac{\partial h}{\partial u} (s, \varphi(s), \dot{\varphi}(s)) \quad (1.4)$$

D'après le théorème de la valeur moyenne, il existe $s_0 \in]t, t+\tau[$ tel que $\dot{\varphi}(s_0) = \frac{x_1 - x_0}{\tau}$ et donc d'après [H3'] p vérifie

$$|p(s_0)| \leq h_1(b)$$

La formule $p(s) = p(s_0) + \int_{s_0}^s \frac{\partial h}{\partial x} (\sigma, \varphi(\sigma), \dot{\varphi}(\sigma)) d\sigma$ permet alors d'écrire

$$|p(s)| \leq h_1(b) + \int_t^{t+\tau} \left| \frac{\partial h}{\partial x} (\sigma, \varphi(\sigma), \dot{\varphi}(\sigma)) \right| d\sigma$$

et d'après H3 , si $(t, \tau, x_0, x_1) \in \Delta_5(t^1)$ on aboutit en posant $C = C(t_1)$ à

$$|p(s)| \leq h_1(b) + C [(M+1) t^1 + I(t, t+\tau, \varphi)]$$

Si $(t, \tau, x_0, x_1) \in \Delta_1(a) \cap \Delta_3(b) \cap \Delta_5(t^1)$ on a donc un d tel que

$$|p(s)| \leq d \quad \forall s \in [t, t+\tau]$$

(4) et (H4') donnent alors, pour tout s de $[t, t+\tau]$

$$|\dot{\varphi}(s)| \leq L_1(d)$$

3. Le problème à un pas.

On suppose H1),...,H4) vérifiées, f s.c.s. de \mathbb{R} dans \mathbb{R} , $\sup f < \infty$. On note pour $t_0 < t_1$

$$Tf(t_0, t_1, x) = \sup_{\substack{\varphi \in C^1[t_0, t_1] \\ \varphi(t_1) = x \\ f(\varphi(t_0)) \geq 0}} f(\varphi(t_0)) - I(t_0, t_1, \varphi)$$

et $\Phi^f(t_0, t_1, x)$ l'ensemble des chemins optimaux c'est-à-dire, des φ de $C^1([t_0, t_1])$ vérifiant $\varphi(t_1) = x$, $f(\varphi(t_0)) \geq 0$ et $f(\varphi(t_0)) - I(t_0, t_1, \varphi) = Tf(t_0, t_1, x)$.

Ce paragraphe est destiné à préparer le suivant (itération). Pour majorer uniformément $\|\dot{\varphi}\|_\infty$ pour $\varphi \in \Phi^f(t, t+\tau, x)$ avec x vérifiant " $Tf(t, t+\tau, x)$ minoré par une constante", nous distinguerons suivant que φ sature ou non la contrainte. Dans le cas de non-saturation nous utilisons les équations d'Euler. Il ne servirait à rien de supposer la fonction initiale f de classe C^1 car de toute façon, les futures fonctions initiales (Tf) seront "seulement" s.c.s. et Lip.loc. Nous rappelons le formalisme des sous différentiels ou gradients généralisés ([3],[4]) puis nous établissons la Lip.loc. de $Tf(t, t+\tau, \cdot)$. A cause de la contrainte, l'argument classique de modification "en parallèle" d'un chemin optimal est inopérant. Nous utiliserons dans ce cas une modification "conique" (respectant la contrainte).

Lemme 2 : (sous les hypothèses H1...H4).

Pour tout (a, t^1) de $\mathbb{R}_+ \times \mathbb{R}_+^*$ il existe a' de \mathbb{R}_+ tel que pour tout f , si t, τ vérifient $t+\tau \leq t^1$, si x vérifie $Tf(t, t+\tau, x) \geq -a\tau$, tout φ de $\Phi^f(t, t+\tau, x)$ vérifiant $\int_0^1 (\varphi(t)) = 0$ satisfait à

$$\sup_{t \leq s \leq t+\tau} |\dot{\varphi}(s)| \leq a' \tag{1.5}$$

Preuve du lemme 2 :

$Tf(t, t+\tau, x) \geq -a\tau$ et $f(\varphi(t)) = 0$ entraîne $(t, \tau, \varphi(t), x) \in \Delta_2(a)$ (notations du lemme 1) ; il existe donc $b > 0$ tel que $(t, \tau, \varphi(t), x)$ appartienne à $\Delta_3(b)$. Or il appartient à $\Delta_5(t^1)$ par hypothèse, donc finalement à $\Delta_1(at^1) \cap \Delta_3(b) \cap \Delta_5(t^1)$ d'où le résultat. \square

On se propose maintenant d'obtenir le même résultat quand $f(\varphi(t)) > 0$.

Définitions : [3][4].

a) Une fonction f de \mathbb{R}^n dans \mathbb{R} est dite localement lipschitzienne si pour tout sous-ensemble borné B de \mathbb{R}^n il existe $K_B > 0$ tel que

$$x_1, x_2 \in B \Rightarrow |f(x_1) - f(x_2)| \leq K_B \|x_1 - x_2\|$$

Une telle fonction est dérivable p.p., la fonction f' ainsi définie est bornée sur les sous-ensembles bornés de son domaine de définition.

b) On appelle gradient généralisé de f en x et on note $\partial f(x)$ l'enveloppe convexe de l'ensemble des limites de la forme $\lim f'(x+h_i)$ où $h_i \rightarrow 0$ quand $i \rightarrow +\infty$.

$\partial f(x)$ est un compact convexe non vide ; dans le cas $n=1$ c'est un intervalle compact. On notera $\sup |\partial f(x)| = \sup \{|a| : a \in \partial f(x)\}$. Si $x \in \overset{\circ}{B}$ on a évidemment $\sup |\partial f(x)| \leq K_B$.

Si f est convexe on retrouve le sous-différentiel, si f est de classe C^1 au voisinage de x on a $\partial f(x) = \{f'(x)\}$.

c) Soit $t_0 < t_1$, $\varphi \in A[t_0, t_1]$. On dit que h de $[t_0, t_1] \times \mathbb{R}^2$ dans \mathbb{R} est lipschitzienne en (x, v) près de φ , s'il existe $\eta > 0$ et $k \in L^1[t_0, t_1]$, tels que l'on ait pour presque tout t de $[t_0, t_1]$:

$$\sup_{i=1,2} [|x_i - \varphi(t)| + |v_i - \dot{\varphi}(t)|] < \eta \Rightarrow |h(t, x_1, v_1) - h(t, x_2, v_2)| \leq k(t) [|x_1 - x_2| + |v_1 - v_2|]$$

Théorème [4]:

Si g est une fonction de \mathbb{R} dans \mathbb{R} , localement lipschitzienne, si h est mesurable en t , φ optimale pour le problème (**) et h lipschitzienne en (x,v) près de φ , il existe $p \in A[t_0, t_1]$ tel que

$$(\dot{p}(t), p(t)) \in \partial h(t, \varphi(t), \dot{\varphi}(t)) \quad \text{p.p. sur } [t_0, t_1] \quad (1.6)$$

$$- p(t_0) \in \partial g(\varphi(t_0)) \quad \square$$

Les outils sont maintenant en place pour démontrer le résultat annoncé, avec la définition suivante que notre contrainte nous amène à proposer.

Définition :

Si A est un sous-ensemble de \mathbb{R} d'intérieur $\overset{\circ}{A} \neq \emptyset$ f est dite localement uniformément lipschitzienne sur A (en abrégé LUL sur A) s'il existe η et $K > 0$ tels que

$$x_1, x_2 \in \overset{\circ}{A}, \quad |x_1 - x_2| < \eta \Rightarrow |f(x_1) - f(x_2)| < K |x_1 - x_2|$$

Lemme 3 : (sous les hypothèses H1...H4)

Si f est LUL sur $\{f \geq 0\}$ et $\sup f < \infty$, alors pour tout (a, t^1) de $\mathbb{R}_+ \times \mathbb{R}_+^*$ il existe a'' de \mathbb{R}_+ tel que $t + \tau \leq t^1$, $Tf(t, t + \tau, x) \geq -a\tau$, $\varphi \in \Phi^f(t, t + \tau, x)$, $f(\varphi(t)) > 0$ entraînent : $\sup_{t \leq s \leq t + \tau} |\dot{\varphi}(s)| \leq a''$.

Preuve du lemme 3 :

Puisque φ ne sature pas la contrainte, on peut toujours modifier f en une fonction localement lipschitzienne sur \mathbb{R} , identique à f dans un voisinage de $\varphi(0)$ et telle que φ reste optimal pour le nouveau problème. Comme φ est de classe C^1 , l'image de $[t, t + \tau]$ par l'application $(\varphi, \dot{\varphi})$ est bornée ce qui joint à l'hypothèse H3 (h est de classe C^1) entraîne que h est lipschitzienne en (x,v) près de φ . Le théorème de Clarke nous donne alors :

$$\dot{p}(s) = \frac{\partial h}{\partial x}(s, \varphi(s), \dot{\varphi}(s)) \quad , \quad p(s) = \frac{\partial h}{\partial v}(s, \varphi(s), \dot{\varphi}(s)) \quad \text{p.p. sur } [t, t+\tau] \quad (1.7)$$

$$- p(t) \in \partial f(\varphi(t))$$

d'où pour tout $s \in [t, t+\tau]$

$$p(s) \in - \partial f(\varphi(t)) + \int_0^s \frac{\partial h}{\partial x}(\sigma, \varphi(\sigma), \dot{\varphi}(\sigma)) \, d\sigma$$

L'hypothèse H3) et l'optimalité de φ donnent alors

$$\sup_{t \leq s \leq t+\tau} |p(s)| \leq a_1 \quad (1.8)$$

avec $a_1 = \sup |\partial f(\varphi(t))| + C [(M+1)t^1 + f(\varphi(t)) - Tf(t, t+\tau, x)]$

et finalement sous les hypothèses du lemme 3 :

$$\sup_{t \leq s \leq t+\tau} |p(s)| \leq \sup |\partial f| + C [(M+1+a)t^1 + (\sup f)^+] = a_2 \quad (1.9)$$

(7) et (H4') permettent alors de conclure

$$\sup_{t \leq s \leq t+\tau} |\dot{\varphi}(s)| \leq a'' = L_1(a_2) \quad (1.10)$$

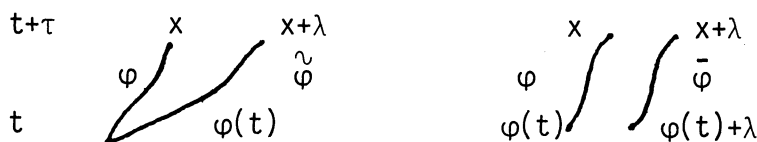
Proposition 2 :

Sous les hypothèses des lemmes 3 et 4, la fonction $f_1 = Tf(t, t+\tau, \cdot)$ est LUL sur $\{f_1 \geq -a\tau\}$.

Preuve de la proposition 2 :

Montrons d'abord la sous-lipschitzianité de f_1 sur \mathbb{R} . Soit x quelconque, $\varphi \in \Phi^f(t, t+\tau, x)$ et $|\lambda| \leq \tau$. Le chemin $\tilde{\varphi}$ défini par $\tilde{\varphi}(s) = \varphi(s) + \lambda \frac{s-t}{\tau}$ (1.11) pour $s \in [t, t+\tau]$ vérifie $\tilde{\varphi}(t) = \varphi(t)$, $\tilde{\varphi}(t+\tau) = x+\lambda$, satisfait toujours la contrainte $f(\tilde{\varphi}(t)) \geq 0$ ce qui n'est pas toujours le cas du chemin classique $\bar{\varphi}$ défini par

$$\bar{\varphi}(s) = \varphi(s) + \lambda \quad (1.12)$$



La sous-optimalité de $\tilde{\varphi}$, jointe à l'optimalité de φ donne :

$$f_1(x+\lambda) - f_1(x) \geq I(t, t+\tau, \varphi) - I(t, t+\tau, \tilde{\varphi})$$

et d'après H3) on a pour tout (ξ, v) , $s \leq t^1$, $|\delta_1| \vee |\delta_2| \leq 1$,

$$h(s, \xi, v) - h(s, \xi+\delta_1, v+\delta_2) \geq -D (|\delta_1| + |\delta_2|) [M+1+h(s, \xi, v)] \quad (1.13)$$

où $D = \exp 2C(t^1)$

ce qui donne, en tenant compte de H2) et en intégrant

$$f_1(x+\lambda) - f_1(x) \geq -D (1+t^1) [M+1+\bar{h}(\|\dot{\varphi}\|_\infty)] |\lambda| \text{ pourvu que } |\lambda| \leq 1 \wedge \tau \quad (1.14)$$

D'après les lemmes 2) et 3), si $f_1(x) \geq -a\tau$ et $\varphi \in \Phi^f(t, t+\tau, x)$ on a

$\|\dot{\varphi}\|_\infty \leq a' \vee a''$ d'où en posant

$$D' = D(1+t^1) [M+1+\bar{h}(a' \vee a'')] \quad (1.15)$$

$f_1(x+\lambda) - f_1(x) \geq -D' |\lambda|$ dès que $f_1(x) \geq -a\tau$ et $|\lambda| \leq 1 \wedge \tau$

on obtient le résultat demandé en échangeant x et $x+\lambda$.

Remarque :

Le majorant B' de $\sup |\partial T f|$ sur $T f \geq -a\tau$ dépend via a'' du majorant de $\sup |\partial f|$ et $\sup f$ (cf. (9) et (10)).

4. Le problème à n pas.

Notre objectif, rappelons-le, est d'étudier l'itération de l'opérateur T . Avec l'idée d'approcher le problème d'optimisation à contrainte permanente le pas de temps τ est destiné à tendre vers 0 ; nous cherchons pour un temps fini t (c'est-à-dire au bout de $O(\frac{1}{\tau})$ opérations) une majoration uniforme des pentes des chemins optimaux et des gradients des fonctions successives.

D'après la remarque du 3), une itération mécanique des majorations conduirait à $+\infty$. Nous prouverons notre résultat par un traitement local, s'appuyant sur les propriétés de dérivabilité p.p., conséquence de la proposition 2. Construisons d'abord notre itération :

f est LUL et vérifie $\sup f < \infty$, τ est positif et ζ_n , $n \in \mathbb{N}$ est une suite de réels.

$$\begin{aligned} \text{Soit} \quad & g_0 = f \\ & \dots \dots \dots \\ n \geq 1 \quad & g_n(x) = \tau g_{n-1}((n-1)\tau, n\tau, x) + \zeta_n \quad \text{si } \{g_{n-1} \geq 0\} \neq \emptyset \\ & = -\infty \quad \text{sinon} \end{aligned} \quad (1.16)$$

On note $F_n = \{g_n \geq 0\}$, $n \geq 0$. La suite ζ_n est là pour des raisons techniques dans les calculs des chapitres suivants. En fait la frontière de la zone de présence a tendance à contenir peu d'individus et les ζ_n sont là pour s'en écarter un peu. Pour que l'algorithme ait des chances de converger vers la limite attendue (cf. introduction générale) il faut néanmoins que les ζ_n soient négligeables par rapport aux g_n c'est-à-dire $\tau^{-1} \sup_n |\zeta_n| \xrightarrow{\tau \rightarrow 0} 0$. Deux cas principaux serviront : $\zeta_n \equiv \zeta > 0$ et $\frac{\zeta}{\tau} \rightarrow 0$ et $\zeta_n \equiv -\zeta$ et $\frac{\zeta}{\tau} \rightarrow 0$.

Proposition 3 : (sous H1...H4)

1) Pour tout n , g_n est LUL sur F_n .

2) Pour tout $a \geq 0$, $t^1 > 0$, il existe C_1, \dots, C_4 tels que pour tout $\tau > 0$, toute suite ζ_k , $k \in \mathbb{N}$ telle que $\sup_k \zeta_k \leq a\tau$ et n tel que $n\tau \leq t^1$ on ait

$$i) \quad \|g_n^+\|_\infty \leq C_1 \quad (1.17)$$

$$ii) \quad \forall x \in F_n \quad \sup \{ \|\dot{\varphi}\|_\infty : \varphi \in \Phi^{g_{n-1}((n-1)\tau, n\tau, x)} \} \leq C_2 \quad (1.18)$$

$$\forall x, x' \in F_n \quad |x-x'| \leq \tau \wedge 1 \Rightarrow |g_n(x) - g_n(x')| \leq C_3 |x-x'|$$

$$iii) \quad F_n \subset [-C_4, +C_4] \quad (1.19)$$

Preuve de la proposition 3 :

1) H1) entraîne facilement

$$\|g_n^+\|_\infty \leq \|f^+\|_\infty + (M+a) n\tau \quad (1.20)$$

et la LUL se prouve par récurrence en utilisant la proposition 2.

2) ii) et iii) - Pour prouver l'existence de C_2 , on montrera l'existence d'un majorant uniforme pour le coefficient de sous-lipschitzianité de g_n sur F_n et on utilisera la dérivabilité p.p. de g_n , conséquence de 1) pour montrer qu'il majore aussi les sous-différentiels correspondants.

Soit $n\tau \leq t^1$ et $x \in \overset{\circ}{F}_n$ supposé non vide. On définit simultanément par récurrence sur $k \leq n$ une suite de chemins (optimaux) $\varphi_k \in C^1([(k-1)\tau, k\tau])$ et une suite de points x_k vérifiant

$$x_n = x = \varphi_n(n\tau) \quad , \quad x_k = \varphi_k(k\tau) = \varphi_{k+1}(k\tau) \quad 1 \leq k < n \quad , \quad x_0 = \varphi_1(0) \quad (1.21)$$

$$\varphi_k \in \Phi^{g_{k-1}}((k-1)\tau, k\tau, x_k) \quad 1 \leq k < n$$

(Les Φ sont non vides d'après la proposition 1).

Si k est fixé et λ vérifie $g_{k-1}(x_{k-1} + \lambda) \geq 0$, la sous-optimalité du chemin $\bar{\varphi}_k$ (cf. (12)) jointe à l'optimalité de φ_k donne :

$$g_k(x_k + \lambda) - g_k(x_k) \geq g_{k-1}(x_{k-1} + \lambda) - g_{k-1}(x_{k-1}) + I((k-1)\tau, k\tau, \varphi_k) - I((k-1)\tau, k\tau, \bar{\varphi}_k)$$

Utilisant (13), une intégration et l'optimalité de φ_k on aboutit à :

$$g_k(x_k + \lambda) - g_k(x_k) \geq g_{k-1}(x_{k-1} + \lambda) - g_{k-1}(x_{k-1}) - |\lambda| D[a''' \tau + g_{k-1}(x_{k-1}) - g_k(x_k)] \quad (1.22)$$

où $a''' = [M+1+a]$.

La méthode consiste alors à itérer cette relation pour aboutir à la sous-lipschitzianité de g_n .

1er cas : La contrainte n'est jamais saturée ($g_k(x_k) > 0$ pour tout $k < n$). Par continuité, il existe $r > 0$ tel que, pour tout $k < n$ et $|\lambda| \leq r$, on ait $g_k(x_k + \lambda) > 0$. La formule (22), valable pour tout $k \leq n$ donne par addition :

$$g_n(x + \lambda) - g_n(x) \geq f(x_0 + \lambda) - f(x_0) - |\lambda| D[a''' n\tau + f(x_0) - g_n(x)]$$

La LUL de f entraîne alors l'existence de K et r_0 tels que

$n\tau \leq t^1$, $|\lambda| \leq r_0 \wedge r$ et $g_n(x) \geq 0$ entraînent

$$g_n(x+\lambda) - g_n(x) \geq - |\lambda| [K + D (a''' t^1 + \|f^+\|_\infty)] = - K_1 |\lambda| \quad (1.23)$$

2ème cas : On sature la contrainte au moins une fois. On pose

$j = \sup \{k \leq n : g_{k-1}(x_{k-1}) = 0\}$. Par continuité il existe r tel que $j \leq k < n$ et $|\lambda| \leq r$ entraînent $g_k(x_k + \lambda) > 0$. Par addition de (22) on obtient

$$g_n(x+\lambda) - g_n(x) \geq g_j(x_j + \lambda) - g_j(x_j) - |\lambda| D [a''' (n-j)\tau + g_j(x_j) - g_n(x)]$$

x_j étant l'extrémité de φ_j qui sature la contrainte en x_{j-1} , (5) et (14) donnent

$$g_j(x_j + \lambda) - g_j(x_j) \geq -D (1+t^1) [M+1+\bar{h} (a')] |\lambda| = - K_2 |\lambda| \quad \text{pour } |\lambda| \leq 1 \wedge \tau$$

De (20) et $g_n(x) \geq 0$ on tire finalement pour $|\lambda| \leq 1 \wedge \tau \wedge r$

$$g_n(x+\lambda) - g_n(x) \geq - [K_2 + D((M+a+a''')t^1 + \|f^+\|_\infty)] |\lambda| = - K_3 |\lambda| \quad (1.24)$$

En rapprochant (23) et (24), on a trouvé K_4 (ne dépendant que de f, a, t^1) et r' (dépendant de f, n, τ, x) tel que les conditions $|\lambda| \leq r'$, $n\tau \leq t^1$, $g_n(x) \geq 0$ entraînent :

$$g_n(x+\lambda) - g_n(x) \geq - |\lambda| K_4 \quad (1.25)$$

D'après le 1°, g_n est dérivable p.p. sur $\overset{\circ}{F}_n$. Si $g'_n(x) \neq 0$, en prenant λ du signe opposé à $g'_n(x)$ et en faisant tendre λ vers 0 dans (25) on obtient $|g'_n(x)| \leq K_4$, a fortiori vrai si $g'_n(x) = 0$. D'après la définition du gradient généralité, on a donc, pour tout x de $\overset{\circ}{F}_n$

$$\sup |\partial g_n(x)| \leq K_4 \quad (1.26)$$

Pour la majoration de $\|\dot{\varphi}_n\|_\infty$ il suffit de revenir à (5), (9), (10) pour conclure

* contrainte saturée : $\|\dot{\varphi}_n\|_\infty \leq a'$

* contrainte non saturée :

$$\|p_n\|_\infty \leq K_4 + C [\|f^+\|_\infty + (2M+2a+1)t^1] = K_5 \Rightarrow \|\dot{\varphi}_n\|_\infty \leq L_1 (K_5)$$

d'où l'existence de C_2 , puis de C_3 d'après la formule (14). Ceci termine la démonstration de ii).

iii) F_0 est compact par hypothèse, et de plus avec les notations de ii) on a pour tout $n \geq 1$ et tout x de F_n

$$x = \varphi_1(0) + \sum_{k=1}^n \int_{(k-1)\tau}^{k\tau} \dot{\varphi}_k(s) ds$$

et donc

$$|x| \leq \sup_{y \in F_0} |y| + C_2 t^1$$

5. Résultat annexe :

On note, pour $t_0 < t_1$, x_0, x_1 quelconques, (sous les hypothèses H1...H4)

$$\bar{I}(t_0, t_1, x_0, x_1) = \inf \{ I(t_0, t_1, \varphi) : \varphi(t_0) = x_0, \varphi(t_1) = x_1 \}$$

$$|\Phi|(t_0, t_1, x_0, x_1) = \sup \{ \|\dot{\varphi}\|_\infty : \varphi \in \Phi(t_0, t_1, x_0, x_1) \}$$

Nous aurons besoin, dans l'étude du modèle, du résultat annexe suivant :

Proposition 4 :

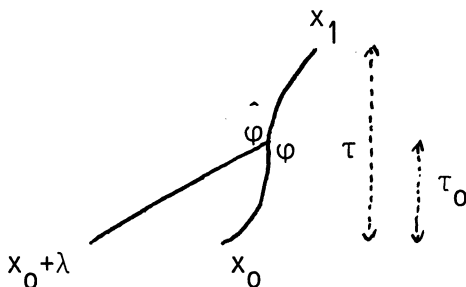
Pour tous $\gamma, d, t^1 > 0$, il existe $\delta_1(\gamma, d)$ et $\delta_2(\gamma, d, t^1) > 0$ tels que pour tous $t, \tau, \lambda, x_0, x_1$ les conditions $|\lambda| \leq d\tau$ et $|\Phi|(t, t+\tau, x_0, x_1) \leq \gamma$ entraînent :

- 1) $\bar{I}(t, t+\tau, x_0+\lambda, x_1) \leq \bar{I}(t, t+\tau, x_0, x_1) + \delta_1 |\lambda|$
- 2) $|\Phi|(t, t+\tau, x_0+\lambda, x_1) \leq \delta_2$, si $t+\tau \leq t^1$

Preuve de la proposition 4 :

Posons $\tau_0 = d^{-1} |\lambda|$ et si $\varphi \in \Phi(t, t+\tau, x_0, x_1)$ soit $\hat{\varphi}$ le chemin défini par

$$\begin{aligned} \hat{\varphi}(s) &= (\varphi(t+\tau_0) - x_0 - \lambda) \frac{s-t}{\tau_0} + x_0 + \lambda & \text{pour } t \leq s \leq t+\tau_0 \\ \hat{\varphi}(s) &= \varphi(s) & \text{pour } t+\tau_0 \leq s \leq t+\tau \end{aligned}$$



La non-optimalité de $\hat{\varphi}$ et l'optimalité de φ donnent :

$$\bar{I}(t, t+\tau, x_0+\lambda, x_1) \leq \bar{I}(t, t+\tau, x_0, x_1) - I(t, t+\tau_0, \varphi) + I(t, t+\tau_0, \hat{\varphi})$$

H1) donne $I(t, t+\tau_0, \varphi) \geq -M\tau_0$ et H2) donne $I(t, t+\tau_0, \hat{\varphi}) \leq \tau_0 \bar{h} \left(\left| \frac{\varphi(t+\tau_0) - x_0 - \lambda}{\tau_0} \right| \right)$

D'après les hypothèses et l'inégalité évidente

$$|\varphi(t+\tau_0) - x_0 - \lambda| \leq |\varphi(t+\tau_0) - \varphi(t)| + |\lambda| \leq (\gamma+d) \tau_0$$

on peut écrire

$$\bar{I}(t, t+\tau, x_0+\lambda, x_1) - \bar{I}(t, t+\tau, x_0, x_1) \leq [M+\bar{h}(\gamma+d)] \tau_0 = d^{-1} [M+\bar{h}(\gamma+d)] |\lambda|$$

Les conditions de l'énoncé placent $(t, t+\tau, x_0+\lambda, x_1)$ dans l'ensemble

$\Delta_1(\gamma+d) \cap \Delta_3(\gamma+d\delta_1 t^1) \cap \Delta_5(t^1)$ et on termine grâce au lemme 1 .

§ II - PRESENTATION DU MODELE :

Il s'agit du modèle de croissance de population introduit dans [6], c'est-à-dire une suite de processus de branchement spatiaux, indexée par un paramètre d'échelle ε , mais avec en plus une population initiale de taille exponentielle réglée par une fonction f_0 . \bar{M} est l'ensemble des mesures de Radon positives sur \mathbb{R} ; M , ensemble des mesures ponctuelles finies (sommées finies de masses de Dirac) est borélien pour la topologie vague. On le munit de la tribu trace notée $\mathcal{B}(M)$.

Soit $U = \sum_{n=0}^{+\infty} \mathbb{N}^n$ et si $u = j_1 \dots j_n \in U$ on note $|u| = n$, $f(u) = j_1 \dots j_{n-1}$. Grâce à [5], on peut construire (cf. [6] chap. I) un espace probabilisable (Ω, \mathcal{I}) d'arbres ω marqués par M (la marque v_u en un noeud u de ω représentant les positions des enfants de u). Pour $u \in U$ on définit le sous-ensemble Ω_u de Ω des arbres admettant u comme noeud et sur Ω_u les v.a. v_u (v_ϕ est noté v), θ_u (arbre translaté en u), $\sigma(u)$ (position associée au noeud u) et $s(u) = (\sigma \circ f^{n-1}(u), \dots, \sigma \circ f(u), \sigma(u))$. Pour tout n , on note $z_n(\omega)$ l'ensemble des noeuds de ω de longueur n , F_n la tribu engendrée par les Ω_u tels que $|u| \leq n$ et les v_u tels que $|u| < n$, et

$$\zeta_n = \sum_{u \in z_n} \delta_{\sigma(u)} \quad Z_n = \sum_{u \in z_n} \delta_{s(u)}$$

Soit $P_{t,x}, (t,x) \in \mathbb{R}_+ \times \mathbb{R}$ une famille mesurable de lois sur $(M, \mathcal{B}(M))$. Pour tout (ε, x) de $\mathbb{R}_+^* \times \mathbb{R}$ on note $S_{\varepsilon,x}$ l'application de $K(\mathbb{R})$ dans lui-même définie par $S_{\varepsilon,x} g(z) = g(x + \varepsilon z)$ et pour tout k de \mathbb{N} , $\bar{p}_{k,x}^\varepsilon$ l'image de $P_{k\varepsilon,x}$ par $S_{\varepsilon,x}$. Pour tout ε , la proposition 1 de [6] permet de construire une famille $P_{k,x}^\varepsilon, (k,x) \in \mathbb{N} \times \mathbb{R}$ de lois sur (Ω, \mathcal{I}) telle que l'image de $P_{k,x}^\varepsilon$ par v soit $\bar{p}_{k,x}^\varepsilon$ et pour tout n , conditionnellement à F_n , les v.a. $\theta_u (u \in z_n)$ sont indépendantes et de lois respectives :

$P_{\bar{\sigma}}^{\varepsilon}(u)$ où $\bar{\sigma}$ est définie par $\bar{\sigma}(\phi) = (k, x)$, $\bar{\sigma}(u) = (k + |u|, \sigma(u))$ si $u \neq \phi$

On suppose que pour tout (t, x) , $P_{t,x}$ a une espérance notée $\mu_{t,x}$ et que pour tout (θ, v)

$$L(t, x, \theta) = \text{Log} \int e^{\theta \xi} \mu_{t,x}(d\xi) \quad , \quad h(t, x, v) = \sup_{\theta} \theta v - L(t, x, \theta) \quad (2.1)$$

sont finies (donc l'enveloppe convexe du support de $\mu_{t,x}$ est \mathbb{R}). $\bar{P}_{k,x}^{\varepsilon}$ a alors une espérance notée $Q_{\varepsilon}(k, x, \cdot)$ dont la log-Laplace L_{ε} et la transformée de Cramer H_{ε} vérifient

$$L_{\varepsilon}(k, x, \theta) = L(k\varepsilon, x, \varepsilon\theta) \quad , \quad H_{\varepsilon}(k, x, v) = h(k\varepsilon, x, \frac{v}{\varepsilon}) \quad (2.2)$$

On suppose données une fonction f_0 , LUL sur son domaine compact de positivité F_0 (cf. I), deux fonctions de \mathbb{R}^+ dans \mathbb{R}^+ , r_0 et s_1 vérifiant quand ε tend vers 0

$$\lim r_0(\varepsilon) = \lim s_1(\varepsilon) = 0 \quad \text{et} \quad \lim \varepsilon^{-1} r_0(\varepsilon) = \lim \varepsilon^{-1} s_1(\varepsilon) = +\infty \quad (2.3)$$

et une famille π_0^{ε} d'éléments de M , à support dans F_0 , telle que pour tout ε et $x \in F_0$

$$|\varepsilon \text{Log} \pi_0^{\varepsilon}(\Delta_{r_0(\varepsilon)}(x)) - f_0(x)| \leq s_1(\varepsilon) \quad (2.4)$$

où pour r positif $\Delta_r(x)$ désigne l'intervalle $[x-r, x+r]$

Tous les éléments sont maintenant en place pour la construction de la famille de processus de branchements spatiaux.

Soit $\hat{\Omega} = \sum_{n=1}^{\infty} \Omega^n$ l'ensemble des réunions finies d'arbres spatiaux, muni de la tribu \hat{I} traçant $I^{\mathbb{N}^n}$ sur Ω^n . Toutes les v.a. définies sur Ω s'étendent de manière naturelle à $\hat{\Omega}$, on les notera avec le signe $\hat{\cdot}$. Pour tout ε on note P^{ε} la probabilité sur $\hat{\Omega}, \hat{I}$ correspondant à la superposition d'arbres spatiaux aléatoires indépendants, ayant pour positions initiales les points chargés par π_0^{ε} et de lois les $P_{0,x}^{\varepsilon}$ correspondants. Dans ces conditions, pour tout n , la v.a. $\hat{\zeta}_n^{\varepsilon}$ est une somme de v.a. indépendantes du type ζ_n .

Enfin on note $\hat{\zeta}_t^{\varepsilon} = \hat{\zeta}_{\lfloor \frac{t}{\varepsilon} \rfloor}$.

§ III - RESULTATS ASYMPTOTIQUES SUR L'ESPERANCE DU PROCESSUS.

On montre ici que la limite de $\varepsilon \text{Log } E^\varepsilon \hat{\zeta}_t^\varepsilon (\Delta_r(x))$ est $Tf_0(0,t,x)$. La minoration en 1 est une adaptation, au cas de population initiale, de la proposition de [6]. La majoration en 2 est plus complexe. De façon générale on cherche à majorer des $E_{\zeta_n}(A)$ avec A borélien de \mathbb{R} . Grâce à une renormalisation on peut utiliser les techniques de grandes déviations pour les chaînes (sous)-markoviennes. Il est bien connu ([1], [8]) que les ensembles adéquats pour la majoration de la probabilité sont liés aux ensembles de niveau de la fonctionnelle I . Il faut donc inclure $\mathbb{R}^{n-1} \times A$ dans un ensemble de niveau bien choisi. Comme les 2 énoncés classiques ([1], [8]) ne sont pas directement applicables, on établit une proposition (dite technique), voisine, avec une démonstration plus simple faite en appendice. Elle nous servira aussi dans le IV pour trouver, par récurrence une majoration d'espérance pour des généalogies contraintes, clé de la majoration finale de $\varepsilon \text{Log } \hat{\zeta}_t^\varepsilon$ grâce à l'inégalité de Markov.

1. Minoration .

Proposition 5 :

Si h vérifie H1...H4 et

[H5] h est de classe C^2 en v et il existe h_2 décroissante de \mathbb{R}^+ dans \mathbb{R}_*^+ telle que

$$v(t,x) \quad |v| \leq r \Rightarrow \frac{\partial^2 h}{\partial v^2}(t,x,v) \geq h_2(r)$$

alors pour tout $(x,t^1) \in \mathbb{R} \times \mathbb{R}_*^+$, il existe $\sigma > 0$ tel que pour tout $t \leq t^1$

$$\lim_{\substack{\varepsilon \rightarrow 0 \\ r \geq \sigma \sqrt{\varepsilon t}}} \varepsilon \text{Log } E^\varepsilon \hat{\zeta}_t^\varepsilon (\Delta_r(x)) \geq Tf_0(0,t,x) \quad (3.1)$$

Preuve de la proposition 5 :

On part de l'inégalité suivante valable pour tout $x_0 \in F_0$

$$E^\varepsilon \hat{\zeta}_t^\varepsilon (\Delta_r(x)) \geq \int_{\Delta_{r_0}(x_0)} E_{0,y}^\varepsilon (\zeta_{[\frac{t}{\varepsilon}]} (\Delta_r(x))) \pi_0(dy) \quad (3.2)$$

D'après les formules (4.9) et (4.11) p. 23 de [6] on a pour tout γ , l'existence de σ (fixe) et de $\delta(\varepsilon)$ tendant vers 0 avec ε tels que pour tout φ vérifiant $\varphi(t) = x$, $\|\dot{\varphi}\|_\infty \leq \gamma$ et tout $r \geq \sigma \sqrt{t\varepsilon}$, on ait :

$$E_{0,\varphi(0)}^\varepsilon \zeta_{[\frac{t}{\varepsilon}]} (\Delta_r(x)) \geq \frac{1}{2} \exp - \frac{1}{\varepsilon} [I(0,t,\varphi) (1+\delta(\varepsilon))+\delta(\varepsilon)] \quad (3.3)$$

Soit $\varphi^X \in \Phi^0(0,t,x)$ ($\subset \Phi(0,t,\varphi^X(0),x)$) et avec les notations de I. 5, soit $\gamma = |\Phi|(0,t,\varphi^X(0),x)$. Dans ces conditions, d'après la proposition 4, il existe δ_1 et δ_2 telles que les relations $|y-\varphi^X(0)| \leq r_0(\varepsilon) \leq t$, $\varphi \in \Phi(0,t,y,x)$ entraînent

$$I(0,t,\varphi) \leq I(0,t,\varphi^X) + \delta_1 r_0(\varepsilon) \quad \text{et} \quad \|\dot{\varphi}\|_\infty \leq \delta_2$$

En prenant $x_0 = \varphi^X(0)$ dans (3.2) et en appliquant (3.3) aux φ précédents (quitte à remplacer γ par δ_2) on obtient pour $r_0(\varepsilon) \leq t$ et $r \geq \sigma \sqrt{t\varepsilon}$

$$E^\varepsilon \hat{\zeta}_t^\varepsilon (\Delta_r(x)) \geq \frac{1}{2} \pi_0 (\Delta_{r_0}(\varphi^X(0))) \exp - \frac{1}{\varepsilon} \{(I(0,t,\varphi^X) + \delta_1 r_0) (1+\delta) + \delta\}$$

Il suffit alors pour conclure d'appliquer (2.3) et (2.4).

2. Majoration :

Nous établissons, après l'énoncé de la proposition technique, une majoration en temps réel (avec ε) pour un ensemble de niveau, qui sert à obtenir la majoration proposée pour $E^\varepsilon \hat{\zeta}_t^\varepsilon$ et prépare celles du IV.

On rappelle ([6], II) qu'un noyau est une fonction de $\mathbb{N} \times \mathbb{R} \times \mathcal{B}(\mathbb{R})$ dans \mathbb{R}^+ , telle que $Q(i, \cdot, A)$ soit mesurable et $Q(i, x, \cdot)$ soit une mesure finie. Il est dit markovien (resp. sous-markovien) si pour tout (i, x) on a $Q(i, x, \mathbb{R})=1$ (resp. ≤ 1). On définit sur \mathbb{R}^k la mesure $Q_k(i, x, \cdot)$ par

$$Q_k(i, x, g) = \int_{\mathbb{R}^k} g(y_1, \dots, y_k) Q(i, x, dy_1) Q(i+1, y_1, dy_2) \dots Q(i+k-1, y_{k-1}, dy_k) \quad (3.4)$$

Proposition technique :

Soit Q^0 un noyau sous-markovien, et $Q_{i,y}^0$, $(i,y) \in \mathbb{N} \times \mathbb{R}$ la famille correspondante de sous-probabilités sur $\mathbb{R}^{\mathbb{N}}$. On note L^0 sa log-Laplace et H^0 sa duale de Legendre (supposées définies partout). On suppose qu'il existe \bar{L}^0 (qu'on peut supposer paire et positive) telle que pour tout θ de \mathbb{R}

$$\sup_{j \in \mathbb{N}, x \in \mathbb{R}} L^0(j,x,\theta) \leq \bar{L}^0(\theta) \quad (3.5)$$

On note \underline{H}^0 la duale de \bar{L}^0 . Soit $k,q \in \mathbb{N}^*$, $n=kq$, ℓ_0 , $u \in \mathbb{R}$

$$\Xi^0(u) = \left\{ \ell \in \mathbb{R}^n : \sum_{r=0}^{q-1} k H^0(rk, \ell_{rk}, \frac{1}{r} (\ell_{(r+1)k} - \ell_{rk})) \geq u \right\} \quad (3.6)$$

et pour $0 \leq r \leq q-1$, $\gamma \in \mathbb{R}_+^*$

$$B_r = \left\{ \ell \in \mathbb{R}^n : \sup_{rk \leq s \leq (r+1)k} |\ell_s - \ell_{rk}| \leq \gamma \right\} \quad (3.7)$$

Alors, les conditions $u_0 \geq q$, $u \leq u_0$ entraînent :

$$Q_n^0(0, \ell_0, \Xi^0(u) \cap \prod_{r=0}^{q-1} B_r) \leq \exp \left\{ -\frac{u-n\Delta}{1+\Delta} + q \text{Log} \frac{6u_0}{q} \right\} \quad (3.8)$$

$$Q_n^0(0, \ell_0, \bigcup_{r=0}^{q-1} B_r^c) \leq 4q \exp - k \underline{H}^0\left(\frac{\gamma}{2k}\right) \quad (3.9)$$

$$\text{où } \Delta = \sup \left\{ \frac{|H^0(j,y',u) - H^0(i,y,u)|}{1+H^0(i,y,u)} ; |i-j| \leq k, |y-y'| \leq \gamma, u \in \mathbb{R} \right\} \quad (3.10) \quad \square$$

La démonstration de cette proposition est reportée au paragraphe VII. Grâce à l'hypothèse H1, une renormalisation de toutes les espérances permet de ramener notre problème au cas sous-markovien précédent.

Proposition 6 :

Soit $\ell_0 \in \mathbb{R}$, $t_0, t^1, \varepsilon, \eta, \tau \in \mathbb{R}_+^*$, $i_0, k, q, n \in \mathbb{N}^*$ vérifiant $\varepsilon < \eta < \tau < t^1 - t_0$, $t_0 = i_0 \varepsilon$, $\eta = k\varepsilon$, $\tau = q\eta$, $n = kq$. Si $\ell \in \mathbb{R}^n$, on note ℓ^η la ligne brisée ayant pour sommets les points $(t_0 + j\eta, \ell_{kj})$ pour $0 \leq j \leq q$ et si $u \in \mathbb{R}$, on pose

$$\Xi(u) = \{ \ell \in \mathbb{R}^n : I(t_0, t_0 + \tau, \ell^\eta) \geq u \} \quad (3.11)$$

Alors, si h vérifie H1,H2,H3, pour tous K, u_0 de \mathbb{R}_+^* les conditions $\delta = (2K+1)\eta D \leq 1$ (cf. 1.13), $u \leq u_0, \varepsilon \leq M\eta$ entraînent

$$E_{i_0, \ell_0}^\varepsilon Z_n(\Xi(u)) \leq \exp \frac{1}{\varepsilon(1+\delta)^2} \left\{ -u + 3\delta\tau (M+1) + \frac{4\tau\varepsilon}{\eta} \text{Log} \frac{6 \tilde{u}_0 \eta}{\varepsilon\tau} \right\} \quad (3.12)$$

$$+ \exp \frac{1}{\varepsilon(1+\delta)^2} \left\{ 4M\tau - \eta h(K) + 4\varepsilon \text{Log} \frac{4\tau}{\eta} \right\}$$

où $\tilde{u}_0 = u_0 + M\tau$.

Preuve de la proposition 6 :

Fixons $\varepsilon, \eta, \tau, \gamma$ vérifiant $0 < \varepsilon < \eta < \tau$ et $\gamma > 0$. D'après H1 (cf. [6] II, Rem.2) le noyau $Q^0 = e^{-M} ({}_\varepsilon Q)$ est sous-markovien. Comme pour tout $A \in \mathcal{B}(\mathbb{R})$ on a

$$E_{i_0, \ell_0}^\varepsilon Z_n(A) = ({}_\varepsilon Q)_n(i_0, \ell_0, A) = e^{nM} Q_n^0(i_0, \ell_0, A) \quad (3.13)$$

on peut appliquer la proposition technique. Les formules de passage sont (cf. 2.2)

$$L^0(i, \cdot, \theta) = L(i\varepsilon, \cdot, \varepsilon\theta) - M, \quad H^0(i, \cdot, u) = h(i\varepsilon, \cdot, \frac{u}{\varepsilon}) + M, \quad \underline{H}^0(\cdot) = M + \underline{h}(\frac{\cdot}{\varepsilon})$$

D'après H3 (cf. 1.13) on a, si $\gamma + \eta \leq 1$

$$\sup \left\{ \frac{|h(t, y, u) - h(t', y', u)|}{M + 1 + h(t, y, u)} ; |t - t'| \leq \eta, |y - y'| \leq \gamma, u \in \mathbb{R} \right\} \leq D(\gamma + \eta)$$

On suppose désormais $\Delta = D(\gamma + \eta) \leq 1$ (et $i_0 = 0$ pour simplifier). Pour $r \leq q$ l'inégalité $r\eta \leq t < (r+1)\eta$ entraîne $|\ell^\eta(t) - \ell_{rk}| \leq \gamma$ pour $\ell \in B_r$, donc (cf. 1.13)

$$M + h(t, \ell^\eta(t), \dot{\ell}^\eta(t)) \leq \Delta + (1 + \Delta) H^0(r, \ell_{rk}, \frac{1}{k} (\ell_{(r+1)k} - \ell_{rk}))$$

ce qui permet de conclure par intégration :

$$(\Xi(u) \cap \bigcap_{r < q} B_r) \subset (\Xi^0(\frac{u + (M - \Delta)\tau}{\varepsilon(1 + \Delta)}) \cap \bigcap_{r < q} B_r) \quad (3.14)$$

Les conditions de la proposition 6 entraînent

$$\frac{u + (M - \Delta)\tau}{\varepsilon(1 + \Delta)} \leq \frac{\tilde{u}_0}{\varepsilon} \quad \text{et} \quad \frac{\tilde{u}_0}{\varepsilon} \geq \frac{\tau}{\eta}$$

Si on pose $\gamma = 2K\eta$ et $\delta = \Delta$, les formules (3.8), (3.13), (3.14) et $\Delta \leq 1$ permettent d'affirmer

$$E_{0, \ell_0}^\varepsilon Z_n (\Xi(u) \cap_{r < q} B_r) \leq \exp \frac{1}{\varepsilon(1+\delta)^2} \left\{ -u + 3\delta\tau (M+1) + \frac{4\tau\varepsilon}{\eta} \text{Log} \frac{6 \tilde{u}_0 \eta}{\varepsilon\tau} \right\}$$

De même (3.9) et (3.13) entraînent :

$$E_{0, \ell_0}^\varepsilon Z_n (U \cap_{r < q} B_r^c) \leq \exp \frac{1}{\varepsilon(1+\delta)^2} \left\{ 4M\tau - \underline{h}(K) + 4\varepsilon \text{Log} \frac{4\tau}{\eta} \right\}$$

Cette double majoration ne dépend pas de la condition initiale donc, d'après la propriété de Markov, vaut pour t_0 quelconque. Ceci termine la preuve. \square

Passons maintenant à l'énoncé final de ce paragraphe.

Proposition 7 :

Si h vérifie H1, ..., H3 et si $r(\varepsilon)$ tend vers 0 avec ε on a pour tous t, x

$$\overline{\lim}_{\varepsilon \rightarrow 0} \varepsilon \text{Log} E^\varepsilon \hat{\zeta}_t^\varepsilon (\Delta_r(x)) \leq T f_0(0, t, x) \quad (3.15)$$

Preuve de la proposition 7 :

Notons qu'elle servira de modèle pour le paragraphe suivant.

Comme $d_0 = \text{diam } F_0$ est fini il existe un recouvrement de F_0 en intervalles I_j centrés sur $y_j \in F_0$, de diamètre $2r_0(\varepsilon)$, $j=1, \dots, \#_0(\varepsilon)$ avec

$$\#_0(\varepsilon) r_0(\varepsilon) \leq d_0 \quad (3.16)$$

On a évidemment, pour tout $A \in \mathcal{B}(\mathbb{R})$

$$E^\varepsilon \hat{\zeta}_t^\varepsilon(A) \leq \sum_{j \leq \#_0} \pi_0^\varepsilon(I_j) \sup \{E_{0, y}^\varepsilon \zeta_n(A) ; y \in I_j\} \quad (3.17)$$

et pour y quelconque, si $S \geq \sup \{T f_0(0, t, z) : z \in A\}$, avec les notations de la proposition 6 et $t = \tau$ on a :

$$E_{0, y}^\varepsilon \zeta_n(A) \leq E_{0, y}^\varepsilon (Z_n (\Xi(f_0(y) - S))) \quad (3.18)$$

La suite consiste à appliquer (2.4) et la proposition 6.

Pour tout $0 < \delta < 1$ et $y \in F_0$ on a d'après (2.4)

$$\varepsilon \text{Log } \pi_0^\varepsilon (\Delta_{r_0}(y)) \leq f_0(y) + s_1(\varepsilon) \leq \frac{f_0(y) + \gamma_0}{(1+\delta)^2} \quad (3.19)$$

où on a posé

$$\gamma_0 = 4 s_1(\varepsilon) + 3\delta \|f_0^+\|_\infty \quad (3.20)$$

Pour tout $y \in F_0$ on a $f_0(y) - S \leq u_0 = \|f_0^+\|_\infty + |S|$. Si D_0 majore la constante de Lipschitz de f_0 , on peut écrire pour r_0 (donc ε) assez petit et $\varepsilon \leq M\eta$ d'après (3.19) et la proposition 6

$$E^\varepsilon \hat{\zeta}_t^\varepsilon(A) \leq \#_0 \exp \frac{1}{\varepsilon(1+\delta)^2} \left\{ \sup_A T f_0(0, t, \cdot) + D_0 r_0 + \gamma_0 + 3\delta t (M+1) + \frac{4t\varepsilon}{\eta} \text{Log} \frac{6 \hat{u}_0 \eta}{\varepsilon t} \right\} \quad (3.21)$$

$$+ \#_0 \exp \frac{1}{\varepsilon(1+\delta)^2} \left\{ 4Mt - \eta \underline{h}(K) + 4\varepsilon \text{Log} \frac{4t}{\eta} + \|f_0^+\|_\infty + \gamma_0 \right\}$$

x et t sont maintenant fixés. On choisit $A = \Delta_r(x)$, D_1 un majorant de la constante de Lipschitz de $T f_0(0, t, \cdot)$ au voisinage de x (possible d'après une adaptation très simple de la proposition 2) et

$$D_1' = D_1 \vee \|f_0^+\|_\infty \vee |T f_0(0, t, x)| \vee D_0$$

Dans la formule (3.21), on peut remplacer $\sup_A T f_0(0, t, \cdot)$ par $T f_0(0, t, x) + D_1 r$ ce qui rapproche de la conclusion.

Supposons pour l'instant $r \leq 1$ (ce qui entraîne $\hat{u}_0 \leq 3 D_1' + Mt$). On choisit η en fonction de ε tel que

$$\varepsilon \ll \eta \ll 1 \quad (3.22)$$

puis K tel que $\eta \underline{h}(K) = 3 D_1' + 4 Mt = D_1''$ ce qui entraîne d'après [H2]

$$\delta \ll 1 \ll K \quad (3.23)$$

et $4\varepsilon \text{Log} \frac{4t}{\eta} \leq D_1'$ pour ε assez petit

On a désormais

$$4 Mt - \eta \underline{h}(K) + 4\varepsilon \text{Log} \frac{4t}{\eta} + \|f_0^+\| \leq -D_1' \leq T f_0(0, t, x)$$

et par conséquent

$$E^\varepsilon \zeta_t^\varepsilon(\Delta_r(x)) \leq 2 \#_0 \exp \frac{1}{\varepsilon(1+\delta)^2} \left\{ T f_0(0,t,x) + D_1'(r+r_0) + \gamma_0 + 3(M+1)t\delta + \frac{4t\varepsilon}{\eta} \text{Log} \frac{6D_1''\eta}{\varepsilon t} \right\}$$

Pour conclure, il faut outre (3.22), (3.23) utiliser (3.16) et (2.3) qui nous assure que $\varepsilon \text{Log}(2 \#_0)$ tend vers 0. \square

Les propositions 6 et 7 sont à rapprocher du théorème 2 de [6] qui, avec l'inégalité de Markov, permet d'aboutir à la conclusion suivante.

Corollaire :

Dans le cas d'un seul ancêtre en x_0 ($f_0(x_0) = 0$, $F_0 = \{x_0\}$) et sous les hypothèses du théorème 2 de [6], pour tout (t, x_0, x_1) de $\mathbb{R}^+ \times \mathbb{R}^2$ s'il existe $\varphi \in \Phi(0, t, x_0, x_1)$ vérifiant

$$\inf_{\substack{k, \varepsilon \\ 0 \leq k\varepsilon < t - \varepsilon}} h(k\varepsilon, \varphi(k\varepsilon), \frac{1}{\varepsilon} (\varphi(k+1)\varepsilon - \varphi(k\varepsilon))) > 0 \quad (3.24)$$

(hypothèse de sous-criticité),

il existe σ positif tel que

$$\lim_{\varepsilon \rightarrow 0} \varepsilon \text{Log} P_{0, x_0}^\varepsilon \left(\zeta_{\left[\frac{t}{\varepsilon}\right]}(\Delta_{\sigma\sqrt{\varepsilon}}(x)) > 0 \right) = \lim_{\varepsilon \rightarrow 0} \varepsilon \text{Log} E_{0, x_0}^\varepsilon \left(\zeta_{\left[\frac{t}{\varepsilon}\right]}(\Delta_{\sigma\sqrt{\varepsilon}}(x)) = -\bar{I}(0, t, x_0, x) \right) \quad (3.25)$$

§ IV - MAJORATION DE $\hat{\zeta}_t^\varepsilon$ EN PROBABILITE.

On se propose de préciser, dans le modèle τ -approché, la description asymptotique de $\hat{\zeta}_t^\varepsilon$. On caractérise ici la zone "désertique" $(F_{\lfloor \frac{t}{\tau} \rfloor}^c)$ et on majore en probabilité $\varepsilon \text{Log } \hat{\zeta}_t^\varepsilon (\Delta_r(x))$ pour $x \in F_{\lfloor \frac{t}{\tau} \rfloor}$ par $g_{\lfloor \frac{t}{\tau} \rfloor}(x)$ (à un facteur $\rho \xrightarrow{\varepsilon \rightarrow 0} 1$ près). Nous utiliserons pour cette démonstration un lemme inspiré de III sur la majoration d'espérances de trajectoires contraintes.

1. Majoration d'espérances contraintes :

τ est un pas de temps et ζ positif. On considère avec f_0 définie au II l'algorithme (cf. 1.16).

$$\begin{aligned}
 &g_0 = f_0 \\
 &\dots\dots\dots \\
 j \geq 1 \quad &g_j = T g_{j-1} ((j-1)\tau, j\tau, \cdot) + \zeta \quad \text{sur } \{g_{j-1} \geq 0\} \neq \emptyset \\
 &= -\infty \text{ sinon.}
 \end{aligned} \tag{4.0}$$

On note toujours $F_j = \{g_j \geq 0\} \quad j \geq 1$

Comme d'habitude $\tau = n\varepsilon$. Posons pour $j \geq 1$ et A borélien de \mathbb{R}

$$\Xi_j^!(A) = \{\ell \in \mathbb{R}^{jn} : \ell_{jn} \in A, \forall 0 < i \leq j-1, \ell_{in} \in F_i\} \tag{4.1}$$

et $N_j(A) = \hat{Z}_{jn}^\varepsilon (\Xi_j^!(A))$

Les 2 cas traités seront $A = (F_j)^c$ et $A = \Delta_r(x)$.

(On s'intéresse toujours aux généalogies aboutissant soit près de x si x est dans la zone de présence, soit dans la zone d'absence, mais dans les 2 cas la contrainte concerne les étapes précédentes, qui toutes doivent correspondre à une situation de présence).

Lemme 4 :

Pour tout t^1, ε , il existe $\tau, \delta, \zeta, \gamma$ dépendant de ε vérifiant : (cf. 3.20, pour γ_0)

$$1 \gg \tau \gg \varepsilon \operatorname{Log} \frac{1}{\varepsilon}, \quad \varepsilon \ll \delta \ll \tau, \quad \varepsilon \ll \zeta \ll \tau, \quad \gamma - \gamma_0 \gg \varepsilon \quad (4.2)$$

tels que $\varepsilon \ll r \ll \tau$ entraîne, pour tout j tel que $j\tau \leq t^1$

$$\hat{E}^\varepsilon(F_j^c) \leq \exp \frac{1}{\varepsilon(1+\delta)^2} \{\gamma_0 - j\gamma\} \quad (4.3)$$

$$\hat{E}^\varepsilon(\Delta_r(x) \cap F_j) \leq \exp \frac{1}{\varepsilon(1+\delta)^2} \{g_j(x) + \gamma_0 - j\gamma\} \quad (4.4)$$

Preuve du lemme 4 :

Elle se fait par récurrence en reprenant, à peu de choses près les idées de la proposition 7 et en appliquant les résultats d'uniformité du I. Supposons $\frac{\zeta}{\tau}$ majoré par 1. D'après (1.19) on peut procéder à un recouvrement de F_{j-1} en $\#_{j-1}$ interv. I_i de diamètre $2r$. Avec les notations de II, on obtient en conditionnant par $F_{(j-1)n}$ (voir aussi [6])

$$E^\varepsilon N_j(A) = \sum_{i \leq \#_j} E^\varepsilon \sum_{u \in Z_{(j-1)n}} 1_{\Xi'_{j-1}}(I_i \cap F_{j-1})(s(u)) E_{(j-1)n, \sigma(u)}^\varepsilon [\zeta_n(\Theta_u(\cdot), A) | F_{(j-1)n}] \quad (4.5)$$

Mais $s(u) \in \Xi'_{j-1}(I_i \cap F_{j-1}) \Rightarrow \sigma(u) \in I_i \cap F_{j-1}$ et vu les propriétés des Θ on a

$$E_{(j-1)n, \sigma(u)}^\varepsilon [\zeta_n(\Theta_u(\cdot), A) | F_{(j-1)n}] \leq \sup_{y \in I_i \cap F_{j-1}} E_{(j-1)n, y}^\varepsilon \zeta_n(A) \stackrel{\text{déf}}{=} a_i \quad (4.6)$$

et donc d'après (4.5)

$$E^\varepsilon N_j(A) \leq \sum_{i \leq \#_{j-1}} a_i E^\varepsilon N_{j-1}(I_i \cap F_{j-1}) \quad (4.7)$$

et a_j se majore à l'aide de la proposition 6 comme dans la preuve de la proposition 7.

D'après la proposition 3, soit D_1 majorant (uniforme en j, τ tels que $j\tau \leq t^1$) à la fois $\|g_j^+\|_\infty$ et la constante de Lipschitz de g_j sur F_j .

Pour $y \in F_{j-1}$, si $S \geq \sup_{z \in A} g_j^+(z)$ on a

$$E_{(j-1)n, y}^\varepsilon (\zeta_n(A)) \leq E_{(j-1)n, y}^\varepsilon (\Xi(g_{j-1}(y) - a + \zeta)) \quad (4.8)$$

et pour $y \in I_i$ on a grâce à (1.18)

$$g_{j-1}(y_i) - \sup_{z \in A} g_j^+(z) - D_1 r + \zeta \leq g_{j-1}(y) - a + \zeta \leq D_1 t^1 \stackrel{\text{d\u00e9f}}{=} u_0$$

d'o\u00f9 si (4.4) est vraie \u00e0 l'ordre $j-1$, d'apr\u00e8s la proposition 6 on a

$$a_i E^\varepsilon N_{j-1}(I_i \cap F_{j-1}) \leq e_1 + e_2$$

$$\text{o\u00f9 } e_1 = \exp \frac{1}{\varepsilon(1+\delta)^2} \left\{ \sup_{x' \in A} g_j^+(x') - \zeta + D_1 r + \gamma_0 - (j-1)\gamma + 3(M+1)\delta\tau + \frac{4\tau\varepsilon}{\eta} \text{Log} \frac{6 \bar{u}_0 \eta}{\varepsilon\tau} \right\} \quad (4.9)$$

$$\text{avec } \bar{u}_0 = u_0 + Mt^1$$

$$\text{et } e_2 = \exp \frac{1}{\varepsilon(1+\delta)^2} \{4M\tau - \eta \underline{h}(K) + 4\varepsilon \text{Log} \frac{4\tau}{\eta} + D_1 + \gamma_0 - (j-1)\gamma\} \quad (4.10)$$

On choisit $\varepsilon \text{Log} \frac{1}{\varepsilon} \ll \eta$, $\eta \underline{h}(K) = 3 D_1 + 4Mt^1$, ε_0 tel que $\varepsilon \leq \varepsilon_0$ entra\u00eene $4\varepsilon \text{Log} \frac{4\tau}{\eta} \leq D_1$, et on peut affirmer que si $\zeta \leq \tau \leq D_1$ on a :

$$4M\tau - \eta \underline{h}(K) + 4\varepsilon \text{Log} \frac{4\tau}{\eta} + D_1 \leq -D_1 \leq \sup_{x' \in A} g_j^+(x') - \zeta \quad (4.11)$$

ce qui entra\u00eene

$$a_i E^\varepsilon N_{j-1}(I_i \cap F_{j-1}) \leq 2 e_1 \quad (4.12)$$

et la majoration est ind\u00e9pendante de i . (4.7) donne alors

$$E^\varepsilon N_j(A) \leq 2 \#_{j-1} e_1 \quad (4.13)$$

et d'apr\u00e8s (1.19), il existe D'_1 tel que $j\tau \leq t^1$ entra\u00eene $\#_j \leq \frac{D'_1}{2r}$, d'o\u00f9

$$E^\varepsilon N_j(A) \leq \frac{D'_1 e_1}{r} \quad (4.14)$$

En notant que si $r \leq \tau \wedge 1$

$$A = \Delta_r(x) \cap F_j \text{ entra\u00eene gr\u00e2ce \u00e0 (1.18) } \sup_{x' \in A} g_j^+(x') \leq g_j(x) + D_1 r$$

$$A = F_j^C \text{ entra\u00eene } \sup_{x' \in A} g_j^+(x') = 0$$

on peut \u00e9crire pour $j \geq 0$

$$(E^\varepsilon N_j(F_j^C)) \vee \left\{ \left(\exp - \frac{1}{\varepsilon(1+\delta)^2} g_j(x) \right) E^\varepsilon N_j(\Delta_r(x) \cap F_j) \right\} \leq \exp \frac{1}{\varepsilon(1+\delta)^2} (\gamma_0 - j\gamma) \quad (4.15)$$

$$\text{o\u00f9 } \gamma = \zeta - \zeta' \quad , \quad \zeta' = 2 D_1 (r \vee r_0) + 3(M+1)\delta\tau + \frac{4\tau\varepsilon}{\eta} \text{Log} \frac{6 \bar{u}_0 \eta}{\varepsilon\tau} + 4\varepsilon \text{Log} \frac{D'_1}{r \wedge r_0} \quad (4.16)$$

(car pour $j=0$ notre hypothèse de récurrence est vraie grâce à (3.20))

Il suffit alors de choisir $1 \gg \tau \gg s_1(\varepsilon) \vee K\eta$ et $\zeta = \zeta' + \gamma_0 + \delta$, $\gamma = \gamma_0 + \delta$ (4.17)

ce qui donne

$$\left. \begin{array}{l} \zeta' \ll \tau \text{ grâce à (4.2) , } \varepsilon \ll r \ll \tau, \text{ et (4.10)} \\ \gamma_0 \ll \tau \text{ grâce à (3.20) et (4.17)} \end{array} \right\} \Rightarrow \zeta \ll \tau$$

$$\delta \ll \tau \text{ grâce à (4.17)}$$

$$\varepsilon \ll \gamma - \gamma_0 = \delta \text{ grâce à (4.10)}$$

Remarque :

Le découpage de $[0, \tau]$ en intervalle de taille η rend le calcul long mais semble incontournable car avec la méthode utilisée, prendre $\tau = \eta$ empêcherait $\frac{\gamma_0}{\tau}$ d'être borné et donc $\frac{\zeta}{\tau}$ de l'être. ■

2. Résultat principal : majoration en probabilité :

Proposition 8 :

Pour tout t, ε , il existe τ, ζ, ρ dépendant de ε , vérifiant

$$1 \gg \tau \gg \varepsilon \text{ Log } \frac{1}{\varepsilon}, \quad \varepsilon \ll \zeta \ll \tau, \quad \varepsilon \ll 1 - \rho \ll \tau \quad (4.18)$$

tels que $\varepsilon \ll r \ll \tau$ entraîne

$$\lim_{\varepsilon \rightarrow 0} \sup_{\substack{t < t^1 \\ t \in \tau \mathbb{N}}} P^\varepsilon \left(\hat{\zeta}_t^\varepsilon \left(F_{\left[\frac{t}{\tau} \right]}^c \right) \geq 1 \right) = 0 \quad (4.19)$$

$$\lim_{\varepsilon \rightarrow 0} \sup_{\substack{t < t^1, t \in \tau \mathbb{N} \\ x \in F_{\left[\frac{t}{\tau} \right]}}} P^\varepsilon \left(\varepsilon \text{ Log } \hat{\zeta}_t^\varepsilon \left(\Delta_r(x) \right) \geq \rho g_{\left[\frac{t}{\tau} \right]}(x) \right) = 0 \quad (4.20)$$

Preuve de la proposition 8 :

Posons $\left[\frac{t}{\tau} \right] = k$, $\tau = n\varepsilon$

On a facilement

$$\{\hat{\zeta}_t^\varepsilon(F_k^C) \geq 1\} \subset \bigcup_{j=1}^k \{N_j(F_j) \geq 1\}$$

(Les généalogies des individus de F_k^C ont quitté les zones de présence - les F_j - pour la première fois au plus tard à l'instant $k-1$). Appliquant l'inégalité de Markov, il vient

$$P^\varepsilon(\hat{\zeta}_t^\varepsilon(F_k) \geq 1) \leq \sum_{j=1}^k E^\varepsilon N_j(F_j)$$

Prenant $\tau, \delta, \zeta, \gamma$ comme dans le lemme et posant $\rho^{-1} = (1+\delta)^2$ on a grâce au lemme :

$$P^\varepsilon(\hat{\zeta}_t^\varepsilon(F_k) \geq 1) \leq \frac{e^{-\rho \frac{(\gamma - \gamma_0)}}{\varepsilon}}{1 - e^{-\rho \gamma / \varepsilon}}$$

De même on peut écrire, si $a_k(x) = \exp \frac{1}{\varepsilon} \rho g_k(x)$

$$\{\hat{\zeta}_t^\varepsilon(\Delta_r(x) \cap F_k) \geq a_k(x)\} \subset \{N_k(\Delta_r(x) \cap F_k) \geq a_k(x)\} \cup \bigcup_{j=1}^{k-1} \{N_j(F_j^C) \geq 1\}$$

et toujours grâce à l'inégalité de Markov :

$$P^\varepsilon(\hat{\zeta}_t^\varepsilon(\Delta_r(x) \cap F_k) \geq a_k(x)) \leq \frac{1}{a_k(x)} E^\varepsilon N_k(\Delta_r(x) \cap F_k) + \sum_{j=1}^{k-1} E^\varepsilon N_j(F_j^C)$$

et toujours grâce au lemme :

$$P^\varepsilon(\hat{\zeta}_t^\varepsilon(\Delta_r(x) \cap F_k) \geq a_k(x)) \leq \frac{e^{-\rho \frac{(\gamma - \gamma_0)}}{\varepsilon}}{1 - e^{-\rho \gamma / \varepsilon}}$$

On termine avec (4.17) . \square

Remarque finale :

Comme il a été précisé dans l'introduction, il reste à étudier la minoration en probabilité de $\varepsilon \text{Log } \hat{\zeta}_t^\varepsilon$ dans la zone de présence et à montrer la convergence de l'algorithme vers une limite, c'est-à-dire remplacer la contrainte séquentielle par une contrainte permanente. Ce sera fait dans nos prochains articles.

§ V - APPENDICE : PREUVE DE LA PROPOSITION TECHNIQUE.

Enonçons tout d'abord un lemme qui servira deux fois dans la démonstration de la proposition.

Lemme T :

Soit $k \geq 1$, $x_0 \in \mathbb{R}$ et B borélien de \mathbb{R}^k . On suppose qu'il existe ℓ^0 de \mathbb{R}^2 dans \mathbb{R} telle que pour tout $\theta \in \mathbb{R}$, la condition $(x_1, \dots, x_k) \in B$ entraîne $G^0(s, x_s, \theta) \leq \ell^0(x_0, \theta)$ pour tout s tel que $0 \leq s \leq k-1$ (T1)

Alors pour tout $v \in \mathbb{R}$ on a : (cf. 3.4)

$$Q_k^0(0, x_0, B \cap \{x_k \geq x_0 + kv\}) \leq \exp - \{k \sup_{\lambda \geq 0} [\lambda v - \ell^0(x_0, \lambda)]\} \quad (T2)$$

$$Q_k^0(0, x_0, B \cap \{x_k \leq x_0 + kv\}) \leq \exp - \{k \sup_{\lambda \leq 0} [+ \lambda v - \ell^0(x_0, \lambda)]\} \quad (T3) \quad \blacksquare$$

Démonstration reportée en fin de paragraphe.

Démonstration de la proposition technique : On suppose $u > 0$ le cas $u \leq 0$ est évident.

L'esprit de la méthode est classique ([1], [8]). On a séparé les trajectoires réalisant $\Xi(u)$ en 2 groupes selon que sur les tranches de durée k elles ont ou non effectué des excursions d'au plus γ .

1) Majoration de $Q_1 = Q_{0, \ell_0}^0 \left(\bigcap_{r=0}^{q-1} B_r \cap \Xi^0(u) \right)$. Posons $\ell_0 = x$ et X la chaîne sous-markovienne associée.

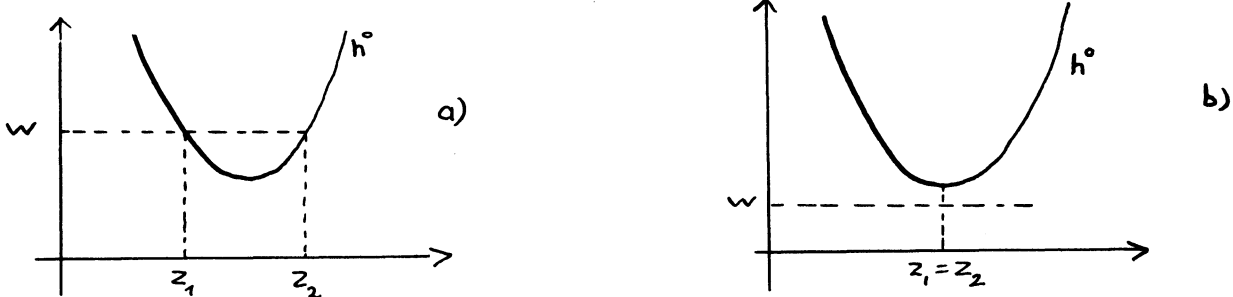
Majorons tout d'abord pour v de \mathbb{R}_+^*

$$p_0 = Q_{0, x}^0 \left(B_0 \cap \left\{ k H^0 \left(0, x, \frac{X_k - x}{k} \right) \geq v \right\} \right) \quad (T10)$$

Si h^0 est la transformée de Cramer d'une sous-probabilité sur \mathbb{R} , l'ensemble $\{z : h^0(z) \geq w\}$ pour $w \in \mathbb{R}$ est la réunion de deux demi-droites $]-\infty, z_1]$, $[z_2, +\infty[$ avec $z_1 \leq z_2$ et

a) si $z_1 < z_2$ alors $h^0(z_1) = h^0(z_2) = w$ et $(h^0)'(z_1) < 0 < (h^0)'(z_2)$ (T.11)

b) si $z_1 = z_2$ alors $h^0(z_1) = h^0(z_2) \geq w$ et $(h^0)'(z_1) = (h^0)'(z_2) = 0$



Ici $h^0 = H^0(0, x, \cdot)$ (x est fixé) et $w = \frac{v}{k}$. On a par conséquent

$$\{k H^0(0, x, \frac{X_k - x}{k}) \geq v\} = \{X_k \leq x + kz_1\} \cup \{X_k \geq x + kz_2\} \quad (T12)$$

On va appliquer le lemme précédent. Pour $j \leq k$ on a pour tout y' tel que $|y' - y| \leq \gamma$ et tout u de \mathbb{R}

$$H^0(0, y, u) \leq H^0(j, y', u) (1 + \Delta) + \Delta \quad (T13)$$

d'où par dualité pour tout θ

$$G^0(j, y', \theta) \leq \ell^0(y, \theta) \quad (T14)$$

où
$$\ell^0(y, \theta) = \frac{1}{1 + \Delta} G^0(0, y, \theta(1 + \Delta)) + \frac{\Delta}{1 + \Delta}$$

On en déduit facilement, grâce à (T2)

$$Q_{0,x}^0(B_0 \cap \{X_k \leq x + kz_1\}) \leq \exp - k \sup_{\lambda \leq 0} \{\lambda z_1 - \ell^0(x, \lambda)\}$$

et

$$\begin{aligned} \sup_{\lambda \leq 0} \{\lambda z_1 - \ell^0(x, \lambda)\} &= \sup_{\lambda \leq 0} \lambda z_1 - \frac{1}{1 + \Delta} G^0(0, x, \lambda(1 + \Delta)) - \frac{\Delta}{1 + \Delta} \\ &= \frac{1}{1 + \Delta} [(\sup_{\theta \leq 0} \{\theta z_1 - G^0(0, x, \theta)\}) - \Delta] \end{aligned}$$

Or d'après (T11)
$$\sup_{\theta \leq 0} \theta z_1 - G^0(0, x, \theta) = \sup_{\theta \in \mathbb{R}} \theta z_1 - G^0(0, x, \theta) = H^0(0, x, z_1) \geq \frac{v}{k}$$

ce qui donne

$$Q_{0,x}^0 (B_0 \cap \{X_k \leq x + kz_1\}) \leq \exp - \frac{v-k\Delta}{1+\Delta} \quad (T15)$$

De même (T3) et (T11) fournissent le même majorant pour $Q_{0,x}^0 (B_0 \cap \{X_k \geq x + kz_2\})$

On en déduit grâce à (T10) et (T12)

$$p_0 \leq 2 \exp - \frac{v-k\Delta}{1+\Delta} \quad (T16)$$

Ce majorant ne dépend pas de la condition initiale et donc si on note

$F_r^1 = \sigma(X_1, \dots, X_{kr})$ on a d'après la propriété de Markov, pour tout x

$$Q_{0,x}^0 [B_r \cap \{k H^0(kr, X_{kr}, \frac{X_{k(r+1)} - X_{kr}}{k} \geq v) \mid F_r^1\}] \leq 2 \exp - \frac{v-k\Delta}{1+\Delta} \quad (T17)$$

Posons pour $0 \leq r \leq q-1$, $Y_{r+1} = k H^0(kr, X_{kr}, \frac{X_{k(r+1)} - X_{kr}}{k}) 1_{B_r}$,

$Y_{r+1} \in \sigma(X_{kr}, \dots, X_{k(r+1)})$. Pour v positif (T17) entraîne

$$Q_{0,x}^0 [Y_{r+1} \geq v \mid F_r^1] \leq \exp - \lambda(v-a) \quad (T18)$$

où $\lambda = \frac{1}{1+\Delta}$ et $a = k\Delta + (1+\Delta) \text{Log } 2 > 0$.

Si v est négatif ou nul la formule (T18) est une évidence.

En posant $Y_r^1 = \lambda(Y_r - a)$, la situation est la suivante : une suite de v.a.

Y_1^1, \dots, Y_q^1 réelles positives, adaptée à la filtration F_1^1, \dots, F_q^1 (on notera F_0^1 la tribu grossière) et vérifiant pour $1 \leq r \leq q$

$$* \quad P^{F_{r-1}^1} (Y_r^1 \geq v) \leq e^{-v} \quad \text{pour tout } v \in \mathbb{R}_*^+$$

le problème étant de majorer $P (Y_1^1 + \dots + Y_r^1 \geq v)$ pour tout $v \in \mathbb{R}_*^+$.

On sait que si V est une v.a. réelle de f.r.F, alors $- \text{Log} [1-F(V)]$

suit une loi exponentielle de paramètre 1. Donc si pour tout r on définit

sur $\mathbb{R} \times \Omega$ Φ_r par $\Phi_r(v, \omega) = - \text{Log} P^{F_{r-1}^1} (Y_r^1 \geq v)$ la v.a. W_r définie

par $W_r(\omega) = \Phi_r(Y_r^1(\omega), \omega)$ est F_r^1 mesurable et suit conditionnellement à F_{r-1}^1

une loi exponentielle, c'est-à-dire que pour tout u de \mathbb{R}_+ on a

$$P^{F_{r-1}^1} (W_r \geq u) = e^{-u}$$

Il en résulte évidemment que les W_r sont indépendants et (*) implique

$$Y'_r \leq W_r \quad \text{pour tout} \quad 1 \leq r \leq q$$

En sommant on obtient

$$P(Y'_1 + \dots + Y'_q \geq v) \leq P(W_1 + \dots + W_q \geq v)$$

Pour majorer cette dernière quantité on utilise l'inégalité de Cramer-Chernov. On rappelle que la transformée de Cramer de la loi exponentielle de paramètre 1 est la fonction qui vaut $+\infty$ sur \mathbb{R}_- et sur \mathbb{R}_+^* :

$$x \rightarrow x - \text{Log } x - 1 = h(x)$$

et donc pour $v \geq q$

$$P(Y'_1 + \dots + Y'_q \geq v) \leq e^{-q h(\frac{v}{q})} = e^{-v+q \text{Log } \frac{v}{q} + q}$$

d'où pour tout $v_0 \geq q$ et tout $v \leq v_0$

$$P(Y'_1 + \dots + Y'_q \geq v) \leq e^{-v + q \text{Log } \frac{v_0}{q}}$$

et en revenant aux Y

$$P(Y_1 + \dots + Y_q \geq v) \leq \exp - \lambda(v-qa) + q + q \text{Log } \frac{v_0}{q}$$

dès que $\lambda v \leq v_0$ et $v_0 \geq q$.

Si on remplace λ et a par leurs valeurs on en déduit facilement

$$Q_1 \leq \exp \left\{ -\frac{1}{1+\Delta} (u-n\Delta) + q \text{Log } \frac{2e u_0}{q} \right\}$$

pourvu que $u \leq u_0$ et $u_0 \geq q$.

2) Majoration de $Q_2 = Q_{0,0}^0 \left(\bigcup_{r=0}^{q-1} B_r^C \right)$. On pose encore $\ell_0 = x$.

On a

$$Q_2 \leq \sum_{r=0}^{q-1} Q_{0,x}^0(B_r^C) \leq q \sup_{r \leq q-1} Q_{0,x}^0(B_r^C) \quad (\text{T19})$$

On majore tout d'abord $Q_{0,x}^0(B_0^C)$. D'après Ventse1 [8], on a

$$Q_{0,x}^0 (B_0^C) \leq 2 \sup_{\substack{0 \leq s \leq s' \leq k \\ y \in \mathbb{R}}} Q_{s,y}^0 \{ |X_{s'} - y| > \frac{\gamma}{2} \} \quad \text{le sup étant pris sur}$$

Posant $\gamma' = \frac{\gamma}{2}$ et appliquant le lemme avec $\ell^0(y, \lambda) = \bar{\Gamma}^0(\lambda)$ pour tout y , on obtient grâce à la parité de $\bar{\Gamma}^0$

$$Q_{s,y}^0 \{ |X_{s'} - y| > \gamma' \} \leq 2 \exp - (s' - s) \underline{H}^0 \left(\frac{\gamma'}{s' - s} \right) \quad (T20)$$

or \underline{H}^0 est positive et croissante sur \mathbb{R}^+ donc on obtient

$$Q_{0,x}^0 (B_0^C) \leq 4 \exp - k \underline{H}^0 \left(\frac{\gamma}{2k} \right)$$

Ce majorant ne dépend pas de la condition initiale et d'après la propriété de Markov vaut donc pour les B_r^C , $1 \leq r \leq q-1$. On conclut donc par

$$Q_2 \leq 4q \exp - k \underline{H}^0 \left(\frac{\gamma}{2k} \right) \quad \blacksquare$$

Démonstration du lemme :

L'identité valable pour tout x, λ de \mathbb{R}^2 et tout $k \geq 1$

$$E_{0,x} \exp \left\{ \lambda(X_k - x) - \sum_{s=0}^{k-1} L^0(s, X_s, \lambda) \right\} = 1 \quad (T21)$$

entraîne pour tout B mesurable par rapport aux X

$$E_{0,x} 1_B \exp \left\{ \lambda(X_k - x) - \sum_{s=0}^{k-1} L^0(s, X_s, \lambda) \right\} \leq 1$$

et donc en appliquant l'hypothèse

$$E_{0,x} 1_B \exp \{ \lambda(X_k - x) - k \ell^0(x, \lambda) \} \leq 1$$

ou encore

$$E_{0,x} 1_B \exp \lambda(X_k - x) \leq e^{k \ell^0(x, \lambda)}$$

qui permet d'obtenir (T2) et (T3) par optimisation sur l'inégalité de Markov pour les moments exponentiels.

Bibliographie :

- [1] AZENCOTT (R.) - RUGET (G.), *Mélanges d'équations différentielles et grands écarts à la loi des grands nombres. Zeits. für Wahr.*, 1977, t. 38, p. 1-54.
- [2] BIGGINS (J.D.), *Chernoff's theorem in the branching random walk. Journ. of Appl. Prob.* 1977, t. 14, p. 630-636.
- [3] CLARKE (F.H.), *Generalized gradients and applications. Trans. of the Amer. Math. Soc.*, 1975, t. 205, p. 247-262.
- [4] CLARKE (F.H.), *The Euler-Lagrange differential inclusion. Journ. of Diff. Eq.*, 1975, t. 19, p. 80-90.
- [5] NEVEU (J.), *Arbres et processus de Galton-Watson (à paraître).*
- [6] ROUAULT (A.), *Probabilités de présence dans un processus de branchement spatial markovien (soumis aux Ann. IHP), voir Chap. F.*
- [7] RUGET (G.), *Large deviations and more or less rare events in population dynamics. Colloque Luminy. Lect. Notes in Biom. n° 49, 1981, p. 388-400.*
- [8] VENTISEL (A.D.), *Rough limit theorems on large deviations for Markov stochastic processes. Theor. of Prob. and its Appl.* 1976, t 21, n° 2, p. 227-242 et t. 21 n° 3, p. 499-512 .

H. Eléments de minoration en zone de présence pour un processus
de branchement spatial markovien.

ELEMENTS DE MINORATION EN ZONE DE PRESENCE POUR UN PROCESSUS DE
BRANCHEMENT SPATIAL MARKOVIEN.

INTRODUCTION :

On poursuit ici l'étude faite en [6] et en [7] du comportement asymptotique d'une population spatiale modélisée par des processus de branchement. Les notations sont celles de [7] § II. Si ε est le paramètre d'échelle - réglant la fréquence des reproductions et la taille des sauts - et f_0 la fonction réglant la taille initiale de la population, la conjecture [8] est que $\hat{\zeta}_t^\varepsilon(\Delta_r(x))$ - population vivant à l'instant t à une distance de x moindre que r - est de l'ordre de $\exp \frac{1}{\varepsilon} \alpha(t,x)$ lorsque $\alpha(t,x) > 0$ et sinon vaut 0 (extinction) avec une probabilité tendant vers 1. α est définie par une formule variationnelle à contrainte permanente :

$$\alpha(t,x) = \sup_{\substack{f_0(\varphi(0)) \geq 0 \\ \varphi(t)=x \\ \forall s \leq t, \tau f_0(0,s,\varphi(s)) \geq 0}} f_0(\varphi(0) - I(0,t,\varphi))$$

où

$$I(t_1, t_2, \varphi) = \int_{t_1}^{t_2} h(s, \varphi(s), \dot{\varphi}(s)) ds$$

$$\tau f_0(0,s,y) = \sup_{\substack{f_0(\varphi(0)) \geq 0 \\ \varphi(s)=y}} f_0(\varphi(0)) - I(0,s,\varphi)$$

τ étant destiné à tendre vers 0, un modèle τ -approché consiste à n'étudier la population qu'aux instants multiples de τ et dans la formule variationnelle à n'imposer la contrainte qu'à ces instants. Dans [7], on a montré, pour τ lié à ε , l'existence d'une zone asymptotiquement "désertique" (avec probabilité tendant

vers 1), et dans son complémentaire la majoration (en probabilité) de $\varepsilon \text{Log } \hat{\zeta}_t^\varepsilon(\Delta_r(x))$ par $\alpha_+^\tau(t,x)$ déterministe correspondant à une contrainte séquentielle. L'objectif suivant consiste donc, dans une zone de présence, à montrer une minoration par un $\alpha_-^\tau(t,x)$ et ultérieurement de montrer la convergence de α_+^τ et α_-^τ vers α . On avance ici dans cette première direction en présentant 2 minorations, chacune correspondant à un empilement (tranches de durée τ) de situations de même "nature" (type "supercritique" ou "sous-critique").

On s'attend à ce que l'essentiel (au sens ε -Log) de la population près de x provienne d'individus dont les lignes généalogiques sont restées dans un tube construit autour d'un chemin aboutissant en x optimal pour le modèle τ -approché. Cette contribution est donc un bon minorant.

La nature foncièrement différente des cas super-critique et sous-critique pour un processus de Galton-Watson se retrouve ici. Si le chemin optimal est resté constamment en zone sous-critique ($\forall s \leq t, h(s, \varphi(s), \dot{\varphi}(s)) > 0$) on utilise l'évaluation déjà faite en [6] pour la probabilité de présence. Si le chemin optimal est resté constamment en zone super-critique ($\forall s \leq t, h(s, \varphi(s), \dot{\varphi}(s)) > 0$) on utilise un argument d'équi-intégrabilité pour la transformation exponentielle de la v.a. Z_n (analogue spatial de $\frac{Z_n}{m^n}$ de Galton-Watson). Il resterait évidemment à montrer qu'on peut passer d'une zone à l'autre. Ce sera l'objet d'un travail ultérieur. La conjecture est étayée par l'idée que malgré leur nature différente, les 2 "types" de comportement donnent des résultats asymptotiquement identiques quand on les regarde en ε -logarithme.

La première partie est consacrée aux hypothèses et à l'énoncé des 2 résultats. La seconde se compose de certaines notations - la plupart déjà vues en [6] - nécessaires aux démonstrations et de l'idée générale sous-jacente à celles-ci. La troisième est la preuve dans le cas "sous-critique permanent". Le cas super-critique est traité dans la quatrième ; il nécessite un lemme dont la preuve constitue la cinquième partie.

§ I - HYPOTHESES ET RESULTATS.

On considère le modèle de croissance de populations décrit dans [7], § II c'est-à-dire une ε -famille de processus de branchement spatiaux avec population initiale. La famille $\mu_{t,x}$, $(t,x) \in \mathbb{R}^+ \times \mathbb{R}$ des espérances des $P_{t,x}$, processus ponctuels de reproduction de référence est supposée vérifier les hypothèses H1), H2), H3), H4, H5) de [7]. On note ν l'élément générique de M et $N = \nu(1)$. Pour tout $(t,x,\psi) \in \mathbb{R}^+ \times \mathbb{R}^2$ on définit $\nu_{t,x}^\psi$ par la formule

$$\frac{d\nu_{t,x}^\psi}{d\nu}(\xi) = \exp \{ \psi \xi - L(t,x,\psi) \} .$$

On considère alors les 3 hypothèses suivantes :

- A1) Pour tout compact K de $\mathbb{R}^+ \times \mathbb{R}$, il existe $c_1 > 0$ tel que
- $$(t,x) \in K \Rightarrow c_1 \leq P_{t,x}(N > 0) .$$
- A2) Pour tout compact K de $\mathbb{R}^+ \times \mathbb{R}$, il existe c_2 et $\delta > 0$ tels que
- $$(t,x) \in K \Rightarrow E_{t,x} N (\text{Log}_+ N)^{1+\delta} \leq c_2$$
- A3) Pour tout compact K de $\mathbb{R}^+ \times \mathbb{R}^2$, il existe c_3 et $\delta > 0$ tels que
- $$(t,x,\psi) \in K \Rightarrow E_{t,x} \nu_{t,x}^\psi(1) [\text{Log}_+ \nu_{t,x}^\psi(1)]^{1+\delta} \leq c_3 .$$

On a les équivalences H1) H2) H3) H4) H5) de [7] \Leftrightarrow H0) H1) H2) H3) H4) H6) H7) de [6]

$$A1) + A3) \quad \Leftrightarrow \quad H5) \text{ de [6]}$$

Définition ([6]) :

On appelle processus ponctuel (sur \mathbb{R}) découplé, (ou de Neymann-Scott [5]) une probabilité P sur $M, \mathcal{B}(M)$ telle que P soit la loi de $\sum_{i=1}^N \delta_{\xi_i}$ où δ est la masse de Dirac, les ξ_i , $i \in \mathbb{N}$ sont indépendantes et de même loi μ^0 et N (à valeurs \mathbb{N}) est indépendante des ξ_i .

Algorithme :

On se donne $\tau, \zeta > 0$ et on définit par récurrence ([7] (1.16))

$$\begin{aligned}
 g_0 &= f_0 \\
 k \geq 1 \quad g_k(x) &= T g_{k-1} \left((k-1)\tau, k\tau, x \right) - \zeta \quad \text{si} \quad (g_{k-1} \geq 0) \neq \phi \\
 &= -\infty \quad \text{sinon}
 \end{aligned} \tag{1.1}$$

On note $F_k = \{g_k \geq 0\}$ pour $k \geq 0$. Pour la majoration on avait besoin de s'écartier un peu de la frontière (reste ζ) pour être sûr d'avoir extinction ici on s'écarte de la frontière de l'autre côté (reste $-\zeta$) pour être sûr d'avoir suffisamment de "pères" à chaque étape.

Les notions de sous-criticité et super-criticité pour des (portions de) chemins se réfèrent aux notations suivantes pour $i < j$ et $c > 0$

$$\underline{T}(c, i, j, \varepsilon) = \{\varphi \in C^1([i\varepsilon, j\varepsilon])\}:$$

$$\forall i \leq k < j \quad h(k\varepsilon, \varphi(k\varepsilon)) , \frac{1}{\varepsilon} [\varphi((k+1)\varepsilon) - \varphi(k\varepsilon)] \geq c\}$$

$$\bar{T}(c, i, j, \varepsilon) = \{\varphi \in C^1([i\varepsilon, j\varepsilon])\}:$$

$$\forall i \leq k < j \quad h(k\varepsilon, \varphi(k\varepsilon)) , \frac{1}{\varepsilon} [\varphi((k+1)\varepsilon) - \varphi(k\varepsilon)] \leq -c\}$$

Dans le cas homogène les chemins optimaux sont des segments de droites. On distingue alors les pentes appartenant à $\{h > 0\}$ (extinction p.s. dans le modèle à un seul ancêtre - d'où leur nom de sous-critiques) et celles appartenant à $\{h < 0\}$ (croissance exponentielle dans le modèle à un seul ancêtre - d'où leur nom de super-critiques) .

La définition de \underline{T} se réfère à l'hypothèse (*) du théorème 1 de [6] , t étant donné on aura besoin d'un "intérieur" de $T(c, 0, \varepsilon^{-1}t, \varepsilon)$, c'est-à-dire en fait : (si $T = \underline{T}$ on définit \underline{S} , si $T = \bar{T}$, on définit \bar{S})

$$\begin{aligned}
 S(r, c, t, \tau, \varepsilon) &= \{\varphi \in T(c, 0, \varepsilon^{-1}t, \varepsilon) : \exists c' : \forall k \leq \tau^{-1}t \\
 &\quad y \in \Delta_r(\varphi((k-1)\tau)) , \psi \in \Phi((k-1)\tau, k\tau, y, \varphi(k\tau)) \\
 &\quad \Rightarrow \psi \in T(c', (k-1)n, kn, \varepsilon)\}
 \end{aligned}$$

On a alors les 2 résultats suivants :

Proposition 1 :

Sous A1, A2, si pour tout t, x le processus $P_{t,x}$ est découplé, alors pour tout $c, t^1, \varepsilon > 0$ il existe $\tau, \zeta > 0$ dépendant de ε , vérifiant $\zeta \ll \tau$, tels que, dès que $\varepsilon \ll r \ll \tau$ on ait

$$\lim_{\varepsilon \rightarrow 0} \sup_{t,x} P^\varepsilon(\varepsilon \text{Log } \hat{\zeta}_t^\varepsilon (\Delta_r(x)) \leq g_{t\tau^{-1}}(x)) = 0$$

où le sup est pris sur l'ensemble des $t \in \tau\mathbb{N} \cap]0, t^1]$ et des $x \in F_{t\tau^{-1}}$ tels qu'un au moins des chemins optimaux aboutissant à x appartienne à $\underline{S}(r, c, t, \tau, \varepsilon)$.

Proposition 2 :

Sous A3, pour tout $c, t^1, \varepsilon > 0$ il existe $\tau, r, \zeta > 0$ et une suite $\eta_k (> 0)$, $k \in \mathbb{N}$ dépendant de ε , vérifiant $\varepsilon \ll r \ll \tau$, $\zeta \ll \tau$, et $\sup_{k \leq t^1} \eta_k \ll 1$ tels que

$$\lim_{\varepsilon \rightarrow 0} \sup_{t,x} P^\varepsilon(\varepsilon \text{Log } \hat{\zeta}_t^\varepsilon (\Delta_r(x)) \leq g_{t\tau^{-1}}(x) + \eta_{t\tau^{-1}}) = 0$$

où le sup est pris sur $t \in \tau\mathbb{N} \cap]0, t^1]$ et $x \in F_{t\tau^{-1}}$ tel qu'un au moins des chemins optimaux aboutissant à x appartienne à $\bar{S}(r, c, t, \tau, \varepsilon)$. On peut choisir τ, ζ comme dans la proposition 1. On notera $\bar{g}_k = g_k + \eta_k$ pour $k \in \mathbb{N}$.

§ II - OUTILS PRINCIPAUX.

On fixe $n = \frac{T}{\varepsilon}$ et on va effectuer dans le modèle exact (non asymptotique) une récurrence sur le nombre de tranches de durée n .

Sur Ω ([7], § II) on a clairement :

$$\zeta_{jn} = \sum_{u \in Z^{(j-1)n}} \zeta_n(\theta_u(\cdot)) \quad (2.1)$$

Si pour $u \in U$, on note $z_n^u = z_n \circ \theta_u$ et $Z_n^u = Z_n \circ \theta_u$, définies sur Ω_u , la proposition 1 de [6] nous dit que conditionnellement à F_k les v.a. Z_n^u ($u \in z_k$) sont indépendantes et de lois ne dépendant que de $\sigma(u)$.

Si $\hat{\cdot}$ se réfère au modèle avec population initiale ([7], § II), alors pour tout couple x, y de réels, $r, r' > 0$ et pour toute famille Λ^W , $w \in \mathbb{R}$ de boréliens de \mathbb{R}^n vérifiant

$$\Lambda^W \subset \mathbb{R}^{n-1} \times \Delta_r(x) \quad (2.2)$$

on a

$$\hat{\zeta}_{jn}(\Delta_r(x)) \geq \sum_{u \in Z^{(j-1)n}} \hat{1}_{\Delta_{r'}(y)}(\sigma(u)) Z_n^u(\Lambda^{\sigma(u)}) \quad (2.3)$$

Sur $\Delta_{r'}(y)$ les $\sigma(u)$ sont voisins et il est raisonnable d'attendre des propriétés uniformes en u pour les $Z_n^u(\Lambda^{\sigma(u)})$. La récurrence tiendra compte de

$$\sum_{u \in Z^{(j-1)n}} \hat{1}_{\Delta_{r'}(y)}(\sigma(u)) = \hat{\zeta}_{(j-1)n}(\Delta_{r'}(y)) \quad (2.4)$$

On se donne r, r_0 une suite x_k , $k \in \mathbb{N}$ de réels et a_k , $k \in \mathbb{N}$ de réels strictement positifs. On pose

$$B_j = \{\hat{\zeta}_{jn}(\Delta_r(x_j)) \leq a_j\} \quad j > 0 \quad B_0 = \{\hat{\zeta}_0(\Delta_{r_0}(x_0)) \leq a_0\} \quad (2.5)$$

La récurrence se construit alors à partir de la formule :

$$P(B_j) \leq P(B_j \cap B_{j-1}^c) + P(B_{j-1}) \quad (2.6)$$

(Dans les 2 preuves, on se donne x à l'instant $j \in \mathbb{N}$ puis on construit

$x_{j-1}, x_{j-2}, \dots, x_0$ suite de pieds de chemins optimaux ([7] (1.21)) pour l'algorithme et $a_k = \exp \frac{1}{\varepsilon} g_k(x_k)$ ou $\exp \frac{1}{\varepsilon} \bar{g}_k(x_k)$.

Le changement exponentiel - introduit pour les noyaux dans [6] § II - va être utilisé pour les v.a. dans IV. Rappelons et étendons cette notion.

Si Q est un noyau, de log-Laplace G et de transformée de Cramer H , on définit pour toute fonction θ de $\mathbb{R}^+ \times \mathbb{R}$ dans \mathbb{R} et $(i, x, \xi) \in \mathbb{N} \times \mathbb{R}^2$

$$\rho_{i,x}^\theta(\xi) = \exp \{ \theta(i,x) (\xi-x) - G(i,x, \theta(i,x)) \} \tag{2.7}$$

$$Q^\theta(i,x, d\xi) = \rho_{i,x}^\theta(\xi) Q(i,x, d\xi) \tag{2.8}$$

et pour $j > i$ et $\ell \in \mathbb{R}^{\mathbb{N}}$

$$R_{i,j}^\theta(\ell) = \prod_{k=i}^{j-1} \rho_{k, \ell_k}^\theta(\ell_{k+1}) \tag{2.9}$$

Fixons n . Pour tout $(i,x) \in \mathbb{N} \times \mathbb{R}$, à $\ell \in \mathbb{R}^{\mathbb{N}}$ on peut associer dans $\mathbb{R}^{\mathbb{N}}$:

$$\ell^{i,x} = (0, \dots, 0, x, \ell_1, \dots, \ell_n, 0, \dots, 0) \tag{2.10}$$

On pose alors

$$R_n^\theta(i,x, \ell) = R_{i, i+n}^\theta(\bar{\ell}) \tag{2.11}$$

Sur Ω_u , on définit la mesure aléatoire $Z_n^{u, \theta}$ sur $\mathbb{R}^{\mathbb{N}}$ par la formule

$$Z_n^{u, \theta}(d\ell) = R_n^\theta(\sigma(u), \ell) Z_n^u(d\ell) \tag{2.12}$$

Elle a le support fini de Z_n^u , mais en général n'est plus une mesure de comptage.

Compte tenu de [6] (1.4) on peut écrire pour toute γ mesurable positive de $\mathbb{R}^{\mathbb{N}}$ dans \mathbb{R}

$$Z_n^{u, \theta}(\gamma) = Z_n(\gamma R_n^\theta(\sigma(u))) = \sum_{v \in Z_{n-1}^u} R_{n-1}^\theta(\sigma(u), s(v)) \nu_v(\gamma_{s(v)} \rho_{\sigma(v)}^\theta) \tag{2.13}$$

On posera
$$W_v = \nu_v(\rho_{\sigma(v)}^\theta) \tag{2.14}$$

On a facilement pour tout (k,y)

$$E_{k,y} Z_n^{u, \theta}(\gamma) = Q_n^\theta(s(u), \gamma) \tag{2.15}$$

Puisque Q^θ est markovien, pour tout t , la constante 1 est fonction propre pour l'opérateur - agissant sur $C_b(\mathbb{R})$ -

$$g \rightarrow Q^\theta(t, \cdot, g)$$

donc $Z_n^{u, \theta}(1)$, $n \in \mathbb{N}$ est une martingale analogue à $(Z_n m^{-n})$, $n \in \mathbb{N}$ de Galton-Watson, et à la martingale du modèle homogène de Biggins ([2], [3]). Les ensembles Λ^W évoqués au début de ce paragraphe seront des tubes généralisés dont nous allons maintenant rappeler les définitions (cf. [6], § II).

Soit $\lambda \in \mathbb{R}^{\mathbb{N}}$ et θ la fonction de $\mathbb{N} \times \mathbb{R}$ dans \mathbb{R} définie par :

$$\theta(k, y) = \frac{\partial H}{\partial v}(k, y, \lambda_{k+1} - \lambda_k) \quad (2.16)$$

On note pour $\ell \in \mathbb{N}$

$$i \in \mathbb{N} : U_i^\lambda(\ell) = \ell_i - \lambda_i$$

$$i < j : V^\lambda(i, j, \ell) = \theta(i, \lambda_i) \{\ell_{i+1} - \ell_i - \lambda_{i+1} + \lambda_i\} + \sum_{k=i+1}^{j-1} \theta(k, \ell_k) \{\ell_{k+1} - \ell_k - \lambda_{k+1} + \lambda_k\} \quad (2.17)$$

et pour $\underline{u}, \underline{v} > 0$

$$\Gamma^\lambda(i, j, \underline{u}, \underline{v}) = \prod_{k=i+1}^j \{ |U_k^\lambda| \leq \underline{u}, |V^\lambda(i, k)| \leq \underline{v} \} \quad (2.18)$$

et (cf. 2.10)

$$\tilde{\Gamma}^\lambda(i, n, \underline{u}, \underline{v}) = \{ \ell \in \mathbb{R}^{\mathbb{N}} : \ell^{i, \lambda_i} \in \Gamma^\lambda(i, i+n, \underline{u}, \underline{v}) \} \quad (2.19)$$

§ III - CHEMIN OPTIMAL RESTANT SOUS-CRITIQUE (preuve de la proposition 1).

Si dans (2.3) on minore $Z_n^u (\Lambda^{\sigma(u)})$ par X_u , son indicatrice de stricte positivité on est ramené à un calcul sur des v.a. de Bernouilli indépendantes.

Lemme 1 :

Si X_1, \dots, X_k sont k v.a. indépendantes de Bernouilli, de paramètres p_1, \dots, p_k on a si $a < p \leq \min p_1, \dots, p_k$:

$$P(S_k \leq ka) \leq \exp - \frac{ka}{2} (\text{Log } \frac{p}{a})^2 \quad (3.1)$$

Ce résultat, dont la démonstration est très facile, permet grâce à (2.6) d'écrire, après avoir remarqué que $x_j \in F_j$ entraîne $a_j = \exp \frac{1}{\varepsilon} g_j(x_j) \geq 1$

$$P(B_j) \leq P(B_{j-1}) + \exp - \frac{1}{2} (\text{Log } \rho_j)^2 \quad j > 1 \quad (3.2)$$

$$1 < \rho_j \leq \frac{a_{j-1}}{a_j} \inf \{P_{(j-1)\tau, y}^\varepsilon (X_u = 1) ; y \in \Delta_{r'}(x_{j-1})\} \quad (3.3)$$

pourvu qu'un tel ρ existe. C'est un problème de probabilités de présence .

Or d'après la proposition 3 de [7] , on peut trouver un majorant universel pour les pentes des chemins φ_j optimaux pour (1.1) cf.[7] (1.21). On construit D_2 comme [6] (4.7) puis

$$\Lambda^W = \overset{\sim}{\Gamma}^\lambda (j-1)n, n, \underline{u}, \underline{v}) \quad (3.4)$$

$$\text{avec } \lambda_k = \varphi(k\varepsilon) , k \in [(j-1)n, jn] \text{ et } \varphi \in \Phi((j-1)\tau, j\tau, w, x_j) \quad (3.5)$$

Le choix $\underline{u} = \varepsilon \underline{v} = 2 \sqrt{D_2 \tau \varepsilon}$ permet d'affirmer grâce à [6] (formules (4.2), (3.10), (4.9), (4.10)), à [7] (proposition 3) et à l'hypothèse A1, A2 que pour tout $t^1, c > 0$, il existe $\mu > 0$ tel que $j\tau \leq t^1$, $y \in \Delta_{r'}(x_{j-1})$, $r' < \tau$

$$P_{(j-1)\tau, y} (X_\phi = 1) \geq \exp - \frac{1}{\varepsilon} (\mu r' + I((j-1)\tau, j\tau, \varphi_j)) \quad (3.6)$$

D'après (1.1) et (3.3) on peut donc choisir

$$\rho_j = \rho = \underset{\text{déf}}{\exp \frac{1}{\varepsilon} \{\zeta - \mu' r\}} \text{ avec } \underline{u} \leq r' = r \vee r_0 \quad (3.7)$$

si l'on impose $\zeta > \mu' r$.

Compte tenu de la population initiale on a aussi :

$$P(B_1) \leq \exp - \frac{1}{2} (\text{Log } \rho_1)^2$$

avec

$$\rho_1 = \exp \frac{1}{\varepsilon} \{f_0(x_0) - s_1(\varepsilon) - g_1(x_1) - I(0, \tau, \varphi_1) - \mu r'\}$$

$$= \exp \frac{1}{\varepsilon} \{\zeta - s_1(\varepsilon) - \mu r'\}$$
(3.8)

Si on pose $\delta = \zeta - s_1(\varepsilon) - \mu r'$, on obtient

$$P(B_j) \leq j \exp - \frac{\delta^2}{2\varepsilon^2} \quad j\tau \leq t^1$$
(3.9)

Pour aboutir il faut maintenant ajuster $r, \delta, \varepsilon, \tau$. Les conditions suivantes sont suffisantes :

$$\tau^{-1} \exp - \frac{\delta^2}{2\varepsilon^2} \ll 1, \quad \varepsilon \ll \delta \ll \tau, \quad r \vee r_0 < \tau$$
(3.10)

Il suffit de remarquer que si $\tau \gg \varepsilon \text{Log } \frac{1}{\varepsilon} + r_0 + s_1(\varepsilon)$ alors $\delta = 2\varepsilon \sqrt{\text{Log } \frac{1}{\varepsilon}}$ et tout r tel que $\tau \geq r \gg \sqrt{\varepsilon\tau}$ conviennent .

§ IV - CHEMIN OPTIMAL RESTANT SUPER-CRITIQUE (preuve de la proposition 2) .

On rappelle qu'à tout chemin φ est associée son ε -discrétisation λ , puis la fonction θ grâce à (2.16). L'entier j est fixé pour l'instant (> 1) si pour $w \in \mathbb{R}$ on choisit $\varphi_w \in \Phi(j-1)_{\tau, j\tau, w, x_j}$ et si Λ^w est une partie de $\Gamma^w((j-1)n, n, \underline{u}, \underline{v})$ avec $\tau \geq r \geq \underline{u}$ alors on déduit de (2.3) et (2.12)

$$\xi_{jn}(\Delta_r(x_j)) \geq \left\{ \sup_{\substack{\ell \in \Lambda^w \\ w \in \Delta_r(x_{j-1})}} R_n^{\theta, w}(\ell) \right\}^{-1} \sum_{u \in \hat{Z}(j-1)_n} 1_{\Delta_r(x_{j-1})}(\sigma(u)) X_u \quad (4.1)$$

où
$$X_u = Z_n^{u, \theta_{\sigma(u)}}(\Lambda^{\sigma(u)})$$

D'après [6] (lemme 1, formules (4.8) et (4.10)) et [7] (propositions 3 et 4) on a facilement l'existence de $\mu > 0$ tel que le sup précédent soit majoré par

$$\exp \frac{1}{\varepsilon} \{ \mu r + I((j-1)_{\tau, j\tau, \varphi_j}) \}$$

L'application de (2.6) se fait sous la forme

$$P(B_j \cap B_{j-1}^c) \leq P_1 \stackrel{\text{déf.}}{=} P\left(\sum_{u \in A} X_u \leq a, I > a_{j-1}\right) \quad (4.2)$$

où $A = \{u \in \hat{Z}(j-1)_n : \sigma(u) \in \Delta_r(x_{j-1})\}$ et $I = \text{card } A$

et

$$a = a_j \exp \frac{1}{\varepsilon} \{ I((j-1)_{\tau, j\tau, \varphi_j}) + \mu r \} = a_{j-1} \exp \frac{1}{\varepsilon} \{ \eta_j - \eta_{j-1} + \mu r - \zeta \}$$

Il va s'agir pour une somme de v.a. $F_{(j-1)_n}$ - conditionnellement indépendantes (les X_u) de grandes déviations, mais du côté où le support est borné ($X_u \geq 0$). La base de l'argumentation sera l'équi-intégrabilité qu'on va assurer en choisissant convenablement la famille Λ^w . On utilise un lemme démontré dans le V (mise bout à bout de tubes de tailles décroissantes).

Lemme 2 :

Pour tout $d, t^1 > 0$, il existe $D > 0$ et $0 < \eta < \frac{1}{2}$ tels que si on note

$$v_p = D 2^{p(\frac{1}{2} + \eta)} \quad u_p = \varepsilon v_p \quad (p \in \mathbb{N}) \quad (4.3)$$

et pour tout $\varphi \in C^1([i\varepsilon, (i+n)\varepsilon])$

$$n = 2^q \quad \Gamma^\lambda(i, n) = \bigcap_{1 \leq p \leq q} \Gamma^\lambda(i+2^{p-1}, i+2^p, u_p, v_p) \quad (4.4)$$

$$n = 2^{q+p} \quad \Gamma^\lambda(i, n) = \Gamma^\lambda(i, 2^q) \cap \Gamma^\lambda(i+2^q, i+n, u_{q+1}, v_{q+1})$$

et enfin $\hat{\Gamma}^\lambda(i, n) = \{\ell \in \mathbb{R}^n : \ell^{i, \lambda_i} \in \Gamma^\lambda(i, n)\}$ (cf. 2.10)

les conditions $\|\dot{\varphi}\|_\infty \leq d$ et $(i+n)\varepsilon \leq t^1$ entraînent :

$$1) \quad E_{i, \lambda_i}^\varepsilon Z_n^\theta(\hat{\Gamma}^\lambda(i, n)) \geq \frac{1}{2} \quad (4.5)$$

2) Sous l'hypothèse A3), pour tout a et $c > 0$, il existe $\delta' > 0$ et c' tels que les conditions $|\lambda_i| \leq a$ et $\varphi \in \mathcal{T}_c(i, n, \varepsilon)$ entraînent

$$E_{i, \lambda_i}^\varepsilon [Z_n^\theta(\hat{\Gamma}^\lambda(i, n)) \text{Log}_+^{\delta'} Z_n^\theta(\hat{\Gamma}^\lambda(i, n))] \leq c' \quad (4.6) \quad \square$$

On est naturellement amené à choisir $\Lambda^W = \hat{\Gamma}^{\lambda^W}(j-1)n, n)$ (dans ces conditions on peut prendre $\underline{u} = D \varepsilon 2^{1/2+n} n^{1/2+n}$). Les hypothèses de la proposition 2 nous assurent alors l'équi-intégrabilité des X_u , $u \in A$. Ceci nous fournit pour tout α de $]0, \frac{1}{2}[$ un β tel que si $X_u^\beta = X_u \wedge \beta$

$$\inf_{u \in A} E [X_u^\beta | \mathcal{F}_{(j-1)n}] > \alpha \quad \text{sur } I > 0 \quad (4.7)$$

Après centrage, (4.2) donne :

$$P^I \left[\sum_{u \in A} (X_u^\beta - E X_u^\beta) \leq a - I\alpha \right] \quad (4.8)$$

Si on choisit $\alpha = \frac{1}{4}$, et si on impose

$$a_{j-1} = 8a \quad (4.9)$$

La formule (4.8) se transforme, grâce à Tchebyshev en :

$$P_1 \leq 64 E \{1_{I > a_{j-1}} (2I - a_{j-1})^{-2} \sum_{u \in A} \text{Var } X_u^\beta\} \quad (4.10)$$

On en déduit facilement :

$$P(B_j \cap B_{j-1}^c) \leq P_1 \leq 64 \frac{\beta^2}{a_{j-1}} \quad (4.11)$$

Au premier pas de la récurrence on a

$P(B_1) \leq P(\sum_{u \in A} X_u \leq a')$ avec $\text{card } A = I$ (déterministe) $\geq a_0 = \exp \frac{1}{\varepsilon} \{f_0(x_0) - s_1(\varepsilon)\}$

d'après [7] Ch. II et $a' = a_1 \exp \frac{1}{\varepsilon} \{\mu r_0 + I(0, \tau, \varphi_1)\}$ d'où

$$P(B_1) \leq 64 \frac{\beta^2}{a_0} \quad (4.12)$$

pourvu que

$$a_0 = 8 a' \quad (4.13)$$

Les conditions (4.9) et (4.13) sont équivalentes à

$$\eta_1 = \zeta - s_1(\varepsilon) - \varepsilon \text{Log } 8 - \mu r_0 \quad \eta_j - \eta_{j-1} = \zeta - \varepsilon \text{Log } 8 - \mu r \quad (j \geq 2) \quad (4.13)$$

D'après la définition de l'algorithme on a toujours $a_k \geq \exp \frac{\eta_k}{\varepsilon}$ ($k \geq 1$) donc d'après (4.11) et (4.12)

$$P(B_j) \leq 64 \beta^2 \left\{ \frac{1}{a_0} + \sum_{j=1}^k \exp - \frac{\eta_k}{\varepsilon} \right\} \quad (4.14)$$

Par hypothèse de la proposition 2, a_0 est minoré par $\exp \frac{1}{\varepsilon} s_1(\varepsilon)$. En prenant

τ et ζ comme à la fin de la preuve de la proposition 1 et

$r \geq \underline{u} = D (2\tau)^{1/2 + \eta} \varepsilon^{1/2 - \eta}$ (r vérifie $\tau \gg r \gg \sqrt{\varepsilon \tau}$) on aboutit à la conclusion demandée.

§ V - PREUVE DU LEMME 2 (EQUIINTEGRABILITE).

L'introduction de tailles décroissantes de sous-tubes a pour objet la preuve de 2). φ est fixé, donc aussi θ et λ .

Pour 1), d'après (2.15) si P^* est la probabilité associée au noyau markovien Q_{i,λ_i}^θ (i est fixé) on doit majorer $P^*(\hat{\Gamma}^\lambda(i,n)^c)$ par $\frac{1}{2}$. On remarque l'inclusion

$$\Gamma^\lambda(i,n) \supset \bigcap_{p:2^{p-1} \leq n} \Gamma^\lambda(i,i+2^p,u_p,v_p) \quad (5.1)$$

qui entraîne évidemment :

$$P^*((\hat{\Gamma}^\lambda(i,n))^c) \leq \sum_{p:2^{p-1} \leq n} P^*((\tilde{\Gamma}^\lambda(i,i+2^p,u_p,v_p))^c) \quad (5.2)$$

Pour tout d, t^1 on sait - évaluations à partir de l'inégalité de Kolmogorov pour les martingales ([4] p. 149 et [6] lemme 2 et formule (4.7)) - qu'il existe $D_2 > 0$ tel que les conditions $\|\dot{\varphi}\|_\infty \leq d$ et $(i+j) \varepsilon \leq t^1$ entraînent pour tout u et $v > 0$:

$$P^*((\tilde{\Gamma}^\lambda(i, i+j, u, v))^c) \geq D_2 j (u^{-2} \varepsilon^2 + v^{-2}) \quad (5.3)$$

si u_p et v_p sont définis par (4.3), les choix $0 < \eta < \frac{1}{2}$ et $4DD_2 \leq 1-4^{-\eta}$ permettent d'aboutir à la conclusion.

Pour 2), on commence par énoncer un lemme d'Asmussen ([1] , p. 23).

Si g est une fonction de \mathbb{R}^+ dans \mathbb{R}^+ croissante et concave (donc sous-additive) \bar{g} la fonction $x \rightarrow x g(x)$ et $Y_i, i \leq p$ des v.a. indépendantes et positives leur somme S_p vérifie

$$E \bar{g}(S_p) \leq \sum_{i=1}^p E \bar{g}(Y_i) + \bar{g}(E S_p) \quad (5.4)$$

On remarque alors ([3]) que la fonction (δ' est fixé positif)

$$\begin{aligned} g(x) &= s_0 x & x < \xi_0 \\ &= s_1 + s_2 \text{Log}^{\delta'} x & x \geq \xi_0 > 1 \end{aligned} \quad (5.5)$$

est concave pour s_0, s_1, s_2, ξ_0 convenablement choisis et que de plus il existe s_3 tel que pour tout x, y on ait

$$g(xy) \leq s_3 (1 + \text{Log}_+^{\delta'} x + \text{Log}_+^{\delta'} y) \quad (5.6)$$

Partons d'un individu situé en x_0 à l'instant $i=0$ pour simplifier l'écriture. (2.12), (2.13) et (2.14) conduisent à la majoration :

$$Z_j^\theta (\hat{\Gamma}^\lambda(0, j)) \leq \sum_{u \in Z_{j-1}} k(u) W_u \quad (5.7)$$

où $k(u) = 1_{\hat{\Gamma}^\lambda(0, j-1)}(s(u)) R_{j-1}^\theta(0, x_0, s(u))$ et les W_u sont F_{j-1} -conditionnellement indépendantes et d'espérance 1.

Une application de (5.4), jointe à la croissance de la fonction \bar{g} et à l'inégalité de Jensen, donne :

$$\begin{aligned} E_{0, x_0} [\bar{g}(Z_j^\theta(\hat{\Gamma}^\lambda(0, j))) | F_{j-1}] &\leq \sum_{u \in Z_{j-1}} k(u) E_{0, x_0} [W_u g(k(u)W_u) | F_{j-1}] \\ &+ \bar{g}(Z_{j-1}^\theta(\hat{\Gamma}^\lambda(0, j-1))) \end{aligned} \quad (5.8)$$

Pour tout j , si e_j est un majorant de $\sup \{R_{j-1}^\theta(0, x_0, \ell) ; \ell \in \hat{\Gamma}^\lambda(0, j-1)\}$ on a

$$\sup \{k(u) ; u \in Z_{j-1}\} \leq e_j$$

D'après la proposition 1 de [6] (propriété de branchement) la loi de W_u ne dépend que de $\sigma(u)$; plus précisément on a ici :

$$E_{0, x_0} [W_u g(k(u)W_u) | F_{j-1}] \leq f_1(\sigma(u)) \quad (5.9)$$

où

$$f_1(y) = e_j^{-1} E_{(j-1)\varepsilon, y} \bar{g}(v_{(j-1)\varepsilon, y}^\theta(1) e_j) \quad y \in \mathbb{R} \quad (5.10)$$

En définissant f_2 de \mathbb{R}^{j-1} dans \mathbb{R} par

$$f_2(\ell_1, \dots, \ell_{j-1}) = f_1(\ell_{j-1}) 1_{\hat{\Gamma}^\lambda(0, j-1)}(\ell_1, \dots, \ell_{j-1}) \quad (5.11)$$

on peut majorer le premier terme du second membre de (5.8) par $Z_{j-1}^\theta(f_2)$, ce qui permet d'écrire la formule de récurrence en déconditionnant (5.8)

$$E_{0,x_0} \bar{g} (Z_j^\theta (\hat{\Gamma}^\lambda(0,j))) - E_{0,x_0} \bar{g} (Z_{j-1}^\theta (\hat{\Gamma}^\lambda(0,j-1))) \leq Q_{j-1}^\theta(0,x_0,f_2) \quad (5.12)$$

On aura terminé quand on aura montré, sous les conditions de l'énoncé que le second membre de (5.12) est le terme général d'une série convergente (uniformément).

Travaillons désormais avec ε . La clé réside dans l'évaluation de e_j . Si $\varphi \in T(c, \rho, j, \varepsilon)$ avec $j\varepsilon \leq t^1$, d'après [6] (lemme 1 et 4.4) et le choix des u_p, v_p on peut trouver D' ne dépendant que de d, t^1, c tel que

$$\sup \{R_{j-1}^\theta(0, x_0, \ell) ; \ell \in \hat{\Gamma}^\lambda(0, j-1)\} \leq \exp \{D' j^{1/2+\eta} - cj\} \quad (5.13)$$

Il existe alors $j_0(D', c)$ tel que

$$j \geq j_0 \Rightarrow D' j^{1/2+\eta} - cj \leq -c \frac{j}{2} \quad (5.14)$$

$$\text{On prend alors } e_j = \exp \left[\{D' j^{1/2+\eta} - cj\} \vee \left(-c \frac{j}{2}\right) \right] \quad (5.14)$$

Dans ces conditions pour $j < j_0$ on a $e_j < \sum_{k=1}^{j_0} e_k = s$ et d'après (5.6) on a

$$f_1(y) \leq s_3 \{1 + \text{Log}_+^{\delta'} s + E_{(j-1)\varepsilon, y} [v_{(j-1)\varepsilon, y}^\theta(1) \text{Log}_+^{\delta'} v_{(j-1)\varepsilon, y}^\theta(1)]\} \quad (5.16)$$

Cette dernière espérance étant majorée d'après A3) pour y appartenant à un compact, on majore ainsi la somme des j_0 premiers termes de 5.12, pourvu que $\delta' \leq 1 + \delta$.

Pour $j \geq j_0$ à partir des évaluations suivantes, faites pour $1 > a > 0$, $\delta' \leq \delta$, $X > 0$, $EX=1$:

$$a^{-1} E \bar{g}(aX) \leq a s_0 E(X^2 1_{aX \leq \xi_0}) + E(X(\text{Log}_+^{\delta'} aX) 1_{aX > \xi_0})$$

$$EX^2 1_{aX \leq \xi_0} \leq e^{1+\delta} \frac{\xi_0}{a} (\text{Log} \frac{\xi_0}{a})^{-1-\delta} EX(\text{Log}_+^{1+\delta} X)$$

$$\text{et } E(X(\text{Log}_+^{\delta'} aX) 1_{aX > \xi_0}) \leq (\text{Log} \frac{\xi_0}{a})^{-1-\delta+\delta'} EX(\text{Log}_+^{1+\delta} X)$$

on déduit, en faisant $a = \exp -c \frac{j}{2}$, $X = v_{(j-1)\varepsilon, y}^\theta(1)$, une majoration du second terme de (5.12) par un $O(j^{-1-\delta+\delta'})$ qui converge donc dès que $\delta' < \delta$.

Remarque :

Dans le cas du modèle découplé (ou "de Neymann-Scott") A2) entraîne A3). En effet, par application de 5.4 on a si g est comme dans 5.5 mais avec β au lieu de δ' et si les indices t, x sont omis :

$$E^N \bar{g} (v^\psi(1)) \leq N E \bar{g} (e^{\psi X - L(\psi)}) + \bar{g} (N)$$

et d'après (5.6)

$$E \bar{g} (v^\psi(1)) \leq s_3 (EN) [1 + (L(\psi)_-)^{\beta} + \exp \{L(\psi(1+\beta)) - L(\psi)\}] + E \bar{g}(N)$$

ce qui permet de conclure.

Bibliographie :

- [1] ASMUSSEN (S.), *Some martingale methods in the limit theory of supercritical branching processes. Adv. in Prob. and Rel. Topics, 1978, vol. 5, p. 1-26 .*
- [2] BIGGINS (J.D.), *Martingale convergence in the branching random walk. Journ. of Appl. Prob., 1977, vol. 14, p. 25-37.*
- [3] BIGGINS (J.D.), *Growth rates in the branching random walk, Zeits. für Wahr., 1979, t. 48, p. 17-34.*
- [4] FREIDLIN (M.I.), VENTSEL (A.D.), *Random Perturbations of Dynamical Systems, Springer, Berlin, 1984.*
- [5] NEYMANN (J.), SCOTT (E.L.), *Processes of Clustering and Applications, in Stochastic Point Processes, edit. P.A.W. Lewis, Wiley, New York, 1972.*
- [6] ROUAULT (A.), *Probabilités de présence dans un processus de branchement spatial markovien (soumis aux Ann. IHP) - Voir Chap. F*
- [7] ROUAULT (A.), *Espérances et majorations pour un processus de branchement spatial markovien (soumis aux Ann. IHP) - Voir Chap. G.*
- [8] RUGET (G.), *Large deviations and more or less rare events in population dynamics. Colloque Luminy. Lect. Notes in Biomat. n° 49 (1981), p. 388-400 .*