UNIVERSITE DE RENNES I
UER de Mathématiques et Informatique
Campus de Beaulieu
35042 - Rennes Cedex

PUBLICATION DES SEMINAIRES
DE MATHEMATIQUES

SEMINAIRE DE PROBABILITES

RENNES 1984

QUELQUES MATERIAUX POUR L'HISTOIRE DE LA THEORIE DES MARTINGALES (1920-1940)

Exposé de Pierre CREPEL le 05.11.1984

.../...

I - INTRODUCTION

Cet exposé, un peu "érudit et technique" n'est qu'une étape pour servir à l'histoire des martingales : l'étape la plus facile pour un mathématicien, celle qui consiste à décrire les aspect explicites du développement de cette notion dans l'entre-deux-guerres. Le travail le plus intéressant reste à faire, on en dira ici seulement quelques mots en remarque.

L'article qui suit est écrit dans des termes familiers aux probabilistes d'aujourd'hui, et nous espérons que, si limité soit-il dans ses ambitions, il pourra leur être utile.

a) Le "corpus classique"

Quand on parle de "martingale" à un mathématicien de ces dernières décennies (éduqué dans l'esprit du livre de J.L. DOOB de 1953, "Stochastic processes"), cela évoque en gros les idées que voici :

Définition : on appelle "martingale" une famille (*) de variables aléatoires (v.a.) (S_t), indexée en général par le temps ($t \in \mathbb{R}$ ou $\mathbb{Z}$), adaptée à une famille croissante de tribus $\mathcal{F}_t$, et telle que

$$E(S_t / \mathcal{F}_s) = S_s \quad \text{p.s.} \quad \text{pour } s \leq t$$

L'exemple le plus classique est celui où $S_n = u_1 + \ldots + u_n$, (u_n) désignant une suite de v.a. indépendantes et centrées.

Mais on pense aussi à d'autres exemples typiques : si S est une v.a. et si $\mathcal{F}_t$ est une famille croissante de tribus, $S_t = E(S / \mathcal{F}_t)$ est une martingale. Il y en a d'autres qui sont devenues classiques, celle qui donne la dérivée d'une mesure, ou bien encore en statistiques le rapport de vraisemblance, etc.

N.B. Toutes les définitions et tous les théorèmes énoncés ici ne sont exprimés ni sous les formes les plus rigoureuses, ni dans leur généralité maximale.

(*) Cette famille est d'ailleurs plutôt notée $x(t)$: l'exposé fera comprendre simplement cette différence de notation significative.

Le théorème d'arrêt

Soient σ et τ deux temps d'arrêt bornés ($\sigma \leq \tau$), alors $E(S_\tau / \mathcal{F}_\sigma) = S_\sigma$ p.s. (avec des définitions connues).

En d'autres termes, le caractère de martingale d'un processus n'est pas affecté par un temps d'arrêt borné, c'est-à-dire par une section aléatoire du temps tenant compte uniquement de la connaissance du passé. Dit encore autrement, pour un jeu équitable, il n'existe pas de stratégie (c'est-à-dire de façon de miser et de quitter le jeu en tenant compte uniquement des coups passés) qui permette de gagner, du moins lorsqu'on a une fortune finie.

Les théorèmes de convergence

Soit (S_n) une martingale à temps discret. Alors, sous diverses hypothèses exprimant qu'elle n'est pas trop grande, S_n converge p.s. C'est le cas, par exemple, si :

- $\sup_n E\, S_n^2 < \infty$, ou même si
- $\sup_n E\, S_n^- < \infty$ (qui inclut notamment le cas de toutes les martingales positives)

Si de plus, la famille (S_n) est équiintégrable, la limite p.s., S_∞, vérifie $E(S_\infty / \mathcal{F}_n) = S_n$ p.s., on dit que la martingale est "fermée".

Les martingales apparaissent comme la notion n° = 1 pour les convergences presque sûres.

Les inégalités

Les théorèmes de convergence reposent en fait sur des inégalités "maximales" dont l'intérêt est d'ailleurs capital en analyse pour l'étude des fonctions harmoniques. En voici un court échantillon. S_n désigne une martingale, λ un nombre positif quelconque.

.../...

- généralisation de l'inégalité de Bienaymé-Tchebycheff :

$$\mathbb{P}\left[\max_{1 \leq k \leq n} |S_k| \geq \lambda\right] \leq \frac{1}{\lambda^2} ES_n^2$$

$$\mathbb{P}\left[\max_{1 \leq k \leq n} S_k \geq \lambda\right] \leq \frac{1}{\lambda}\left[ES_n^- + ES_1\right]$$

- l'inégalité de Doob (*)

Soit $[a,b]$ un intervalle. Notons $N = N(a,b)$ le nombre (aléatoire) de passages en montant de la suite (S_n) par dessus $[a,b]$, c'est-à-dire le nombre d'indices $j = j(\omega)$ tels que

$S_{j-1} < a$ et $S_j > b$, alors :

$$E(N) \leq \frac{1}{b-a} E(|S_n| + |a|)$$

- et bien d'autres...

Martingales à temps continu

Les exemples-types de martingales à temps continu sont le mouvement brownien, le processus de Poisson "recentré". D'autre part, si Z_t est un processus de Markov et si h est une fonction harmonique, alors $h(Z_t)$ est une martingale. Lorsque t tend vers l'infini, certaines propriétés de convergence sont analogues à celles du cas discret.

D'autre part, sous des hypothèses de séparabilité, les trajectoires des martingales à temps continu ont des propriétés de régularité remarquables (existence presque partout de limites à gauche et à droite, ect.) sur lesquels nous n'insisterons pas.

Enfin, les martingales servent de base aux théories de l'intégrale stochastique et des équations différentielles stochastiques.

.../...

(*) En fait Doob a démontré de nombreuses inégalités de martingales et l'expression "inégalité de Doob" ne désigne pas la même chose chez tous les auteurs.

Remarques :

Nous arrêterons ici ce rappel semi-informel de "ce qui traîne spontanément dans la tête du probabiliste de base des années 60 à propos des martingales". Voir par exemple [1] ou [2].

Tout le monde n'a certes pas le même jugement sur l'importance relative des divers théorèmes, certains sont plus portés sur l'aspect convergence presque sûre, d'autres sur les inégalités, d'autres encore sur le "calcul stochastique". En tout cas, dans les années 60, le problème de l'extension du "théorème limite central", et d'autres résultats connexes, aux martingales rencontre fort peu d'écho (il y aura évolution dans les années 70, cf. [3]).

Qu'en était-il en 1930 ?

b) Les martingales dans le "folklore" des mathématiciens vers 1920-1930

La "préhistoire" n'est pas ici notre objectif. Nous nous contenterons de quelques observations tirées de l'exposé de B. Bru [4].

- L'idée qu'il est impossible de trouver un système de jeu permettant de gagner lors d'une succession de parties équitables est très ancienne : B. Bru l'a trouvée même dans l'Antiquité, chez Xénophon !

L'astuce qui consiste, dans un jeu de pile ou face équitable, à doubler sa mise après chaque perte et à s'arrêter à la première victoire est également connue depuis longtemps et, en tout cas, nommée "martingale" au moins depuis le 18e siècle (*) (voir par exemple Condorcet : "Eléments du calcul des probabilités", Royez 1805, p. 119).

Les traités sur les jeux de casinos abondent au 19e siècle et décrivent diverses méthodes analogues, laissant entrendre plus ou moins clairement qu'on ne peut être sûr de gagner ainsi que si l'on dispose d'une fortune infinie.

.../...

(*) Le mot "martingale" lui-même est plus ancien et employé en des sens divers ; son étymologie est controversée (voir les dictionnaires).

- D'un point de vue plus mathématique, tous "les grands noms" du calcul des probabilités ont étudié des problèmes tels que "la durée du jeu", "la ruine du joueur"..., où l'on peut trouver ce que nous appelerions aujourd'hui des techniques ou méthodes de martingale.

Parmi les plus explicites, citons :

- de Moivre : "De Mensura Sortis"
- Ampère : "Considérations sur la théorie mathématique des jeux" (1802) B.N.
- Ch. Babbage : "An examination of some questions connected with games of chance", Trans. Royal Soc. Edinburgh IX (1820), p. 153-177

Lacroix utilise le mot dans son Traité de Calcul des Probabilités.
On pourrait multiplier les citations, où sont formalisées, sur des jeux précis, les idées du caractère inévitable de la ruine du joueur ayant une fortune finie ou du caractère illusoire de la variation des mises. Mais, même si les principes sont énoncés philosophiquement de manière générale, il ne peut y avoir alors d'énoncés mathématiques généraux : rappelons que ne sont dégagées à ces époques ni la notion de variable aléatoire, ni celle de convergence presque sûre, ni celle d'espérance conditionnelle. (Pour mieux comprendre pourquoi, voir [6]).

Au début du 20e siècle, malgré les travaux de Bachelier(*), Markoff, Borel, Cantelli...., le concept de "martingale" en reste à peu près à ce point, dans la culture des probabilistes.

c) L'ambiance probabiliste des années 1920-1930

Ici encore, il ne s'agit que de points de repère pour la compréhension de la suite et non d'une analyse historique.
H. Cramér [5] définit la décennie 1920-1929 comme une décennie de préparation, et 1930-1939 comme celle des "grands changements".
Le "calcul des probabilités" n'est pas toujours considéré alors comme faisant partie intégrante des mathématiques ; ses charmes un peu flous sont diversement appréciés selon les pays, selon les convictions philosophiques, selon les circonstances de leur utilisation [6] : il y a peu de temps encore, la référence (philosophique et mathématique, contestée ou non) c'était Laplace.

.../...

(*) qui mériteraient une étude à part.

Le mot "probabiliste" ne semble guère utilisé, les mathématiciens qui s'adonnent aux probabilités ont en général une longue carrière d'analyste derrière eux.

- Les traditions nationales

En France, le calcul des probabilités, montré du doigt comme scandaleux au 19e siècle par Auguste Comte, Cauchy..., reste marginal, malgré les Traités de Bertrand, Poincaré, Borel ; jusque vers 1930 on ignore jusqu'au nom de Markov ; quant à la statistique mathématique, elle reste dans la vieille tradition.

En Angleterre, au contraire l'école statistique mathématique est florissante et originale depuis Galton, Pearson, Fisher..., mais la théorie des probabilités y intéresse fort peu de monde.

En Allemagne et dans les pays nordiques, l' "école continentale de statistique", occupée par l'actuariat, la démographie, l'astronomie n'a pas procédé au renouvellement britannique ; il n'empêche que ses apports aujourd'hui méprisés ou oubliés ne sont pas négligeables. Pour une large part les probabilités n'y ont de sens qu'au vu de problèmes pratiques : la théorie de Von Mises est déjà en germe dans cette école.

C'est en Russie, puis en U.R.S.S., que la recherche théorique en probabilités a une avance considérable, d'abord avec Tchebychev puis ses élèves Markov, Lyapounov, ensuite avec la jeune génération d'après la Révolution d'Octobre (Kolmogorov, Khintchine...).

On oublie souvent à tort les travaux des Italiens (sauf Cantelli) très en pointe en 1900 et 1920, ainsi que les tentatives "abstraites" en Pologne (Łomnicki, Steinhaus). Quant à la faiblesse des Etats-Unis elle tranche par rapport à ce que sera l'école américaine après la guerre.

.../...

Les grandes préoccupations (explicites) des "probabilistes", du moins conformément aux apparences, jusque vers 1935, voir par exemple [5].

Les théorèmes limites pour sommes de variables aléatoires indépendantes et quelques travaux pionniers pour se dégager petit à petit de cette condition d'indépendance.
Points de repère : loi forte des grands nombres (LFGN) et la loi du logarithme itéré (LLI) : Cantelli (1917), puis Kolmogoroff et Khintchine(1924-29) diverses versions du théorème limite central (TLC) et de sa réciproque : Lindeberg et P. Lévy (1922), Lévy et Feller (1935) ... ; les travaux de Markov à partir de 1906, puis ceux de S.Bernstein à partir de 1917-1918.

Les processus stochastiques : "de Markov", à temps discret puis continu, jugés fondamentaux à partir de la fin des années 20 : voir notamment les articles d'Hostinsky et surtout Kolmogorov (1931) ; le mouvement brownien : Wiener(1923) ; les processus à accroissements indépendants : P. Lévy et Khintchine (1934-1937) ; les processus stationnaires : Khintchine (1934).

Les fondements :discussions philosophiques et mathématiques sur les bases de la théorie des probabilistes sur son champ d' "applications", sur le lien entre probabilité et fréquence, sur l'axiomatique : S. Bernstein (1917), Von Mises (1919, →), Łomnicki et Steinhaus (1923), Kolmogorov (1933), etc.
A noter d'ailleurs que l'apport des "Grundbegriffe" de Kolmogorov n'est pas pour l'essentiel ce qu'en dit la légende : en fait les liens entre probabilités et théorie de la mesure sont familiers depuis longtemps aux probabilistes.
La nouveauté radicale vient plutôt de la place de l'indépendance dans cette axiomatique et dans la théorie des probabilités, de l'espérance conditionnelle, de la possibilité vraiment opératoire d'utiliser les probabilités dans les espaces de fonctions, donc d'une gestion rigoureuse des processus stochastiques, etc. (C'est d'ailleurs ce que Kolmogorov dit lui-même dans l'introduction de ce livre, qui marque certes une coupure dans l'histoire des probabilités).

.../...

Les liens entre probabilités et statistiques:

Avec l'estimation, les tests... le besoin est ressenti de mieux faire le lien entre les théorèmes limites du calcul des probabilités et les méthodes statistiques : notons par exemple dans ce cadre les théorèmes dits de Glivenko- Cantelli et de Kolmogorov-Smirnov.

Bien sûr, les considérations précédentes cumulent tous les défauts des notices historiques faites par les mathématiciens (attention portée exclusivement au développement interne des théories, jugement à l'aune de nos habitudes actuelles, anachronismes dans le langage etc.) ; mais rappelons qu'elles ne sont là que pour faciliter la lecture des probabilistes en fonction de leurs points de vue habituels.

d) "Martingale" : une notion qui "monte" de partout

Continuant sur le même ton, on pourrait dire que dans les années 20 et 30, petit à petit, des morceaux du "corpus classique" sont présents implicitement de plus en plus dans les travaux des analystes et des probabilistes. Notons, sans le détailler, un échantillon de ces "crypto-martingales" :

- la formalisation de l'impossibilité d'un système de jeu consistant à choisir, pile ou face, les coups où l'on mise et ceux où l'on ne mise pas (Von Mises, Wald...) : voir le chapitre IV.

- le sentiment d'une forte analogie, en même temps que de différences, entre les problèmes de convergence des séries de fonctions orthogonales, en particulier des séries de Fourier, et les théorèmes limites pour les séries de variables aléatoires indépendantes et centrées. Les mêmes mathématiciens démontrent des résultats voisins des deux types au moyen, souvent, de ce qu'on appellerait aujourd'hui des inégalités de martingales avant la lettre : ils ont clairement conscience de la nécessité de dégager des "notions intermédiaires" entre suite orthogonale et suite indépendante centrée : c'est très net chez Kolmogorov, P. Lévy, Kač, Steinhaus, Marcinkiewicz, Zygmund, Paley, Wiener, etc.

.../...

Mais ce n'est pas dit comme ça : on ne trouvera guère, même s'ils en sont tout à fait pénétrés, un énoncé du type suivant : "Soit $\Omega = [0,2\pi]$ muni de la probabilité uniforme, alors $\{\sin nx\}$ est une suite de v.a. orthogonales, mais non indépendantes".

- l'étude du développement en fractions continues d'un nombre pris "au hasard" dans $[0,1]$, à partir de motivations diverses, dont les petites perturbations des planètes [7] , [8]; voir le chapitre III.

- les problèmes de dérivation et le besoin d'une théorie rigoureuse de l'ingration dans les espaces de dimension infinie, d'extensions "propres" du théorème de Fubini qui reviennent (cf. chapitre III) à conditionner une v.a. par une suite croissante de tribus, dirions-nous aujourd'hui.

- les études sur le rapport de vraisemblance.

e) Présentation de la suite

Par souci de clarté, et de simplicité, nous centrerons la suite sur les contributions majeures de 4 auteurs intervenus directement dans l'élaboration de la théorie des martingales :

S. Bernstein, P. Lévy, J. Ville, J. L. Doob ; nous nous contenterons de décrire leurs travaux originaux et d'y ajouter quelques commentaires limités sur leurs circonstances.

A peu de choseprès, ces auteurs se connaissent, se citent : les choses sont claires ; il y a quand même quelques "querelles de priorité" (polies), comme disent les historiens des sciences. Par exemple certains travaux de S. Bernstein écrits en russe et non traduits, restent mal connus ou sous-estimés surtout à l'approche et au début de la guerre.

Pour harmoniser les notations, nous désignerons en général par $u_1, \ldots, u_n, \ldots$ une suite de v.a. vérifiant des conditions diverses d'indépendance ou de presque-indépendance et par $S_n = u_1 + \ldots + u_n$ la suite des sommes partielles.

.../...

Autant que possible nous énoncerons les théorèmes principaux à la fois dans le langage de l'auteur et dans un langage moderne.

Nous indiquerons, autant que faire se peut, la date de réception des articles par la revue de parution ainsi que l'auteur du résumé dans Zentralblatt (ZB), à partir de 1931, et dans les Maths. Reviews (MR) à partir de 1940.

II LES ARTICLES DE S. BERNSTEIN

a) Liste des travaux

En raison du caractère difficilement accessible de certains papiers originaux de S.Bernstein, nous donnons ici une traduction de leurs titres et revues de parution, établie à partir du tome IV des Oeuvres (en russe). Ceux qui nous intéressent le plus directement sont numérotés 4, 8, 22, 23, 24, 25, et 27, cités [B4], [B8] etc. dans la suite.

Ils sont d'ailleurs écrits en français, sauf [B25] qui contient cependant un résumé

[B22]	ZB 18	p. 032 (Khintchine)	
[B23]	ZB 22	p. 243 (de Finetti)	
[B24]	ZB 22	p. 061 (Hostinsky)	; MR 1 n° 340 (Doob)
[B25]	ZB 24	p. 263 (Hostinsky)	; MR 2 n° 107 (Doob)
[B27]			; MR 6 n° 88 (Feller)(*)

Articles historiques :

[B] Divers articles sur S.N. Berstein dans "Russian Math Surveys 24 (1969)

(*) A partir de la fin des années 30, Zentralblatt est tombé à peu près directement sous la coupe des nazis, certains de ses animateurs (comme Feller) qui ont d'ailleurs émigré ont fondé les Mathematical Reviews aux Etats-Unis, et l'intérêt des résumés d'ailleurs très irréguliers du Zentralblatt durant la guerre a beaucoup baissé.

.../...

OEUVRES DE S.N. BERNSTEIN
TOME 4 : PROBABILITES, STATISTIQUE MATHEMATIQUE
1911-1946

.../...

.../...