NORTH-HOLLAND
MATHEMATICS STUDIES

72

Probabilities and Potential B

Theory of Martingales

CLAUDE DELLACHERIE

PAUL-ANDRÉ MEYER

NORTH-HOLLAND
MATHEMATICS STUDIES 72

Probabilities and Potential B

Theory of Martingales

CLAUDE DELLACHERIE

Département de Mathématiques
Université de Rouen
France

and

PAUL-ANDRÉ MEYER

Institut de Recherche Mathématique Avancée
Strasbourg, France

Translated and prepared by
J. P. Wilson

1982

NORTH-HOLLAND PUBLISHING COMPANY – AMSTERDAM · NEW YORK · OXFORD

ISBN: 0 444 86526 8

Translation of:
Probabilités et Potentiel B

Publishers:
NORTH-HOLLAND PUBLISHING COMPANY
AMSTERDAM · NEW YORK · OXFORD

Sole distributors for the U.S.A. and Canada:
ELSEVIER SCIENCE PUBLISHING COMPANY, INC.
52 VANDERBILT AVENUE
NEW YORK, N.Y. 10017

PRINTED IN THE NETHERLANDS

FOREWORD

The second part of our book is devoted to martingale theory; it is longer than the first and what remains to be written, to fulfil the promise in our title, is even longer. At any rate the reader will find here plenty of probability and a (somewhat pale) appearance of the word potential.

This volume contains the classical martingale theory, for discrete and continuous time, decomposition theory for supermartingales and several subjects not mentioned in the first edition: local martingales, quasimartingales, semimartingales, the spaces H^1 and BMO and Burkholder's inequalities. We also complete the account of general process theory begun in Chapter IV. We have also included the theory of stochastic integrals, after some hesitation, for this takes us some way from potential theory.

On the other hand, we say nothing about stochastic differential equations and almost nothing about probabilistic applications of martingale and stochastic integral theory (representations, predictability, filtrations,...). On these subjects the reader should consult Jacod's book [4]: Calcul stochastique et problèmes de martingales.

Our presentation doesn't claim to be either definitive (the theory is advancing too quickly) or complete. Nor do we make any pedagogical claims and it would be unreasonable to use our book for teaching without serious pruning - but we have tried to explain clearly what we are discussing and to provide sufficient comments.

We are most grateful to those who pointed out mistakes or possible improvements. In particular: C.S. Chou, C. Doléans-Dade, M. Emery, E. Lenglart, B. Maisonneuve, C. Stricker, K.A. Yen and M. Yor.

C. Dellacherie
P.A. Meyer

COMPLEMENTS TO CHAPTER IV

We were hampered in writing Chapters VI and VII by the fact that certain properties had not been given sufficiently explicitly in Chapter IV. We regroup them here.

We assume given $(\Omega, F, \mathbb{P})$ with a filtration (F_t) satisfying the usual conditions.[1] We also assume that $F_{0-} = F_0$.

A If H is a right closed predictable set, its debut D_H is a predictable time.[2]

Next we need to enumerate the jumps of an adapted cadlag process.

B THEOREM. Let (X_t) be an adapted cadlag real-valued process. We make the convention $X_{0-} = X_0$. Then the random set

$$U = \{(t, \omega) : X_t(\omega) \neq X_{t-}(\omega)\}$$

is the union of a sequence of disjoint graphs $[\![T_n]\!]$ of stopping times. If X is predictable, the T_n can be chosen to be predictable.

Proof. We could use the sledge hammer theorem IV.117 (in the Appendix to Chapter IV). It would be better to use more modest means.

We set $U_n = \{(t, \omega) : |X_t(\omega) - X_{t-}(\omega)| > 2^{-n}\}$ $(n \geq 0)$ and then $V_0 = U_0$, $V_n = U_n \setminus U_{n-1}$ $(n > 0)$; the sets V_n are optional (predictable if X is predictable) and disjoint.

---

[1] See Remark E at the end of these complements.

[2] For a rather more general result, cf. VIII.11. See Vol. 1, nos. 88 B, C, D.

We then set

$$D_n^1(\omega) = \inf\{t : (t, \omega) \in V_n\}, \quad D_n^{k+1}(\omega) = \inf\{t > D_n^k(\omega) : (t, \omega) \in V_n\}$$

so that D_n^i is the i-th jump of X with amplitude lying between 2^{-n} and 2^{-n+1}; since X is cadlag, the set V_n has no accumulation point at a finite distance and the stopping times D_n^i enumerate the points of V_n. It follows from A that the D_n^i are predictable if X is predictable.

It only remains to renumber the double sequence (D_n^i) as a sequence (T_n).

REMARKS. (1) This argument applies equally well to a process with values in a separable metrizable space E: the condition $|X_t - X_{t-}| > 2^{-n}$ would be replaced by $d(X_{t-}, X_t) > 2^{-n}$, where d is a distance defining the topology of E.

(2) See D below for an analogous, but rather more delicate, property.

C THEOREM. Let X be a cadlag real-valued process. Then X is predictable if and only if the following two conditions are satisfied

(1) For every totally inaccessible stopping time T, $X_T = X_{T-}$ a.s. on $\{T < \infty\}$.

(2) For every predictable time T, X_T is F_{T-}-measurable on $\{T < \infty\}$.

Proof. Suppose X is predictable. Then condition (2) is satisfied for every (not necessarily predictable) stopping time T by IV.67. On the other hand, we saw in B above that the set

$$U = \{(t, \omega) : X_t(\omega) \neq X_{t-}(\omega)\}$$

is a countable union of graphs of predictable times and (1) follows immediately.

Conversely, suppose that conditions (1) and (2) hold. We represent the set U as a countable union of graphs of stopping times S_n and decompose each S_n into its totally inaccessible part S_n^i and its accessible part S_n^a (IV.81 (c)). By condition (1) $S_n^i = \infty$ a.s. and $S_n = S_n^a$ is hence accessible. Then the graph of S_n is contained in a union of graphs of predictable times $(S_{nk})_{k\in\mathbb{N}}$ (IV.81 (a)). We set $V = \bigcup_{n,k} [\![S_{nk}]\!]$; by renumbering the double sequence (S_{nk}) as a single sequence, we can represent V as a union of graphs $[\![T_m]\!]$ of predictable times, which we can then easily make disjoint using a construction by induction on m.[1]

For all m, the r.v. X_{T_m} and X_{T_m-} are $\mathcal{F}_{T_m-}$-measurable: the first by condition (2) and the second by IV.67. The same is true of $\Delta X_{T_m} = X_{T_m} - X_{T_m-}$ and by IV.67 there exists a predictable process (Y^m_t) such that $Y^m_{T_m} = \Delta X_{T_m}$ on $\{T_m < \infty\}$. On the other hand, the graph $[\![T_m]\!]$ is predictable. Then denoting by X_- the process $(X_{t-})_{t\geq 0}$, with $X_{0-} = X_0$, which is left continuous and hence predictable, we have

$$X = X_- + \Sigma_m Y^m I_{[\![T_m]\!]}$$

and this shows that X is predictable.

REMARK. The same result is true for processes with values in a separable metrizable space E: it suffices to embed E in $[0, 1]^{\mathbb{N}}$ and apply Theorem C to each coordinate.

We now return to a result similar to Theorem B.

D THEOREM Let (X_t) be a right continuous adapted real-valued process. We make the convention $X_{0-} = X_0$. Then the random variable

$$U = \{(t, \omega) : X_{t-}(\omega) \text{ does not exist or } X_{t-}(\omega) \neq X_t(\omega)\}$$

is the union of a sequence of disjoint graphs $[\![T_n]\!]$ of stopping times.[2] If X is predictable, the T_n can be chosen to be predictable.

Proof. We shall use a less explicit method than that of Theorem B, without using the sledge hammer IV.117. This will illustrate the possibilities offered by Chapter IV. We deal with the predictable case.

---

[1] It suffices to set $A_m = [\![T_m]\!] \setminus \bigcup_{n<m} [\![T_n]\!]$; A_m is the graph of the required stopping time T_m (cf. IV.88).

[2] To within an evanescent set.

We show first that U is predictable. For this we introduce the processes

$$Y_t^+ = \limsup_{s \uparrow\uparrow t} X_s, \quad Y_t^- = \liminf_{s \uparrow\uparrow t} X_s$$

which are predictable by IV.90. Then U is the union of the two predictable sets $\{Y^+ \neq X\}$ and $\{Y^- \neq X\}$.

To complete the proof it then suffices by IV.88 to show that U is contained (to within an evanescent set) in a countable union of graphs of positive random variables. We construct such r.v. - which are in fact stopping times - as follows: let $\varepsilon > 0$; by transfinite induction we set

$$T_0^\varepsilon = 0, \quad T_{\alpha+1}^\varepsilon = \inf\{t > T_\alpha^\varepsilon : |X_t - X_{T_\alpha^\varepsilon}| > \varepsilon\},$$

$$T_\beta^\varepsilon = \sup_{\alpha<\beta} T_\alpha^\varepsilon \quad \text{if } \beta \text{ is a limit ordinal.}$$

As X is right continuous, $T_{\alpha+1}^\varepsilon > T_\alpha^\varepsilon$ on $\{T_\alpha^\varepsilon < \infty\}$ and it follows that there exists a countable ordinal γ_ε from which onwards $T_\alpha^\varepsilon = +\infty$ a.s. (Chapter 0, no. 8). Then U is contained in the union of all the graphs of the T_α^ε, for $\varepsilon = 1/n$ $(n \in \mathbb{N})$ and $\alpha \leq \gamma_\varepsilon$: the instants where the left limit does not exist last appear, for sufficiently small ε, among the T_β^ε corresponding to the limit ordinals and the jump instants among the $T_{\alpha+1}^\varepsilon$.

REMARK. We now indicate that this result extends easily to processes with values in a compact metrizable space and a little less obviously to processes with values in a separable metrizable space E: for the latter case we embed E in a compact metrizable space F and note that

$$U = \{X_{t-} \text{ does not exist in E}\} \cup \{X_{t-} \text{ exists in E and } X_{t-} \neq X_t\}$$

$$= \{X_{t-} \text{ does not exist in F}\} \cup \{X_{t-} \text{ exists in F and } X_{t-} \notin E\}$$

$$\cup \{X_{t-} \text{ exists in E and } X_{t-} \neq X_t\}$$

$$= \{X_{t-} \text{ does not exist in F}\} \cup \{X_{t-} \text{ exists in F and } X_{t-} \neq X_t\}$$

since X_t always belongs to E. This thus reduces to the same problem, but considering X as a process with values in F.

E REMARK. Some of the above properties remain true without any hypothesis on the family (F_t). First of all A: if H is a right closed predictable set, its debut D_H is not necessarily a random variable since the σ-fields are not assumed to be complete. But if D_H is a stopping time of the family (F_t) or only of the family (F_{t+}), then D_H is a predictable time. For then $[\![D_H, \infty[\![= H \cup]\!] D_H, \infty[\![$, a predictable set.

Similarly, if H is a right closed optional set and D_H is a stopping time of (F_{t+}), the set $[\![D_H, \infty[\![$ is optional and D_H is a stopping time of (F_t).

We now pass to B. Instead of arguing as in the text, we adapt the argument of D. Let $\varepsilon > 0$; we construct inductively

$$T_0^\varepsilon = 0, \quad T_{n+1}^\varepsilon = \inf\{t > T_n^\varepsilon : |X_t - X_{T_n^\varepsilon}| \geq \varepsilon \text{ or } |X_{t-} - X_{T_n^\varepsilon}| \geq \varepsilon\}.$$

It was shown in no. IV.64 that these r.v. are stopping times of (F_t) and tend to ∞ as $n \to \infty$. If X is predictable, the T_n^ε are predictable by A. The set U is contained in the union of the graphs $[\![T_n]\!]$ for $n \in \mathbb{N}$, $\varepsilon = 1/m$ $(m \in \mathbb{N})$ and it remains to make a slight modification (replace each T_n^ε by $(T_n^\varepsilon)_{A(n,\varepsilon)}$, where $A(n, \varepsilon)$ is the event "X jumps at instant T_n^ε") to represent U exactly as a countable union of graphs.

We shall not dwell on C and D: in the proof of D, it would be necessary to consider the inf of the $t > T_\alpha^\varepsilon$ such that X_t <u>or a cluster point of</u> $X_\cdot$ <u>at</u> t- is at distance $\geq \varepsilon$ from $X_{T_\alpha^\varepsilon}$.

CONTENTS

CHAPTER V

GENERALITIES AND THE DISCRETE CASE

This chapter contains on the one hand the fundamental theorems of martingale theory (inequalities, the stopping theorem, convergence theorems) in their discrete form, and on the other hand a number of applications which appeared in Chapter VIII of the first edition. This is a "classical" chapter, which appears in much the same form in all books on probability later than Doob's book [1]. We have abstained from substantially modifying it - the notes entitled "Martingales and stochastic integrals" (Meyer [6]) give another version, containing a few additional results. In this book our emphasis is much more on martingale theory in continuous time, developed in Chapters VI and VII.

1. DEFINITIONS AND GENERAL PROPERTIES

It is interesting to set up the notion of a martingale in all its generality. Therefore in the following definition we denote by $\mathbb{T}$ a set with an order relation denoted by $\leq$. We then very quickly restrict ourselves to the case where $\mathbb{T}$ is an interval of the set of integers (the continuous case, where $\mathbb{T}$ is an interval of $\mathbb{R}$, will be studied in later chapters).

The notions of an increasing family of σ-fields and an adapted process (IV.11 and IV.12) extend to arbitrary ordered sets in the obvious way.

1 DEFINITION. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, $(\mathcal{F}_t)_{t \in \mathbb{T}}$ an increasing family of sub-σ-fields of $\mathcal{F}$ and $X = (X_t)_{t \in \mathbb{T}}$ a real-valued process adapted to the family $(\mathcal{F}_t)$. X is called a martingale (resp. supermartingale, submartingale) with respect to the family $(\mathcal{F}_t)$ if

(1) *each random variable* X_t *is integrable*;

(2) *for every ordered pair* (s,t) *of elements of* $\mathbb{T}$ *such that* $s \le t$,

$$(1.1) \qquad \mathbb{E}\,[X_t | \mathcal{F}_s] = X_s \text{ a.s. (resp. } \mathbb{E}\,[X_t | \mathcal{F}_s] \le X_s, \ \ge X_s).$$

2 REMARKS (a) The notion of a martingale - and the word itself - were introduced by Ville in a remarkable work (Ville [1]) to which we shall return in the historical comments. The notion of a submartingale (formerly called a "semimartingale", cf. Doob [1]) was defined and used by Snell [1]. We state once and for all that it was Doob who proved almost all the fundamental results and used them on all the battle-fields of probability theory, in such a way that no probabilist can any longer afford to ignore martingale theory.

(b) A process X is a submartingale if and only if -X is a supermartingale. Therefore we shall restrict ourselves to studying one of the two classes of processes - usually that of *super*martingales, which is more frequently used in potential theory.

(c) A stochastic process X, given without reference to a family of σ-fields, is called a martingale (resp. supermartingale) if it satifies Definition 1 with respect to its natural family of σ-fields $\mathcal{F}_t = \mathcal{T}(X_s, s \le t)$.

(d) Definition 1 has a number of more or less interesting generalizations. The heart of martingale theory consists of results about *real*-valued processes indexed by the *integers* or the *reals*, defined on a *probability* space whose random variables are *integrable* and satisfy (1.1). It is possible to relax one or other of these hypotheses and get "generalized" theories

- the vast theory of *vector-valued martingales*, which we scarcely touch on in this book[1];

- the theory of martingales *whose time set is not "linearly ordered"* (the case where $\mathbb{T} = \mathbb{R}_+^n$ or $\mathbb{N}_+^n$ for example[2]);

- martingale theory *over a σ-finite measured space* about which we shall say a few words (nos. 39-43);

[1] The basic results are given in Neveu [2], Chapter V, §2 (pp. 100-114).

[2] See especially Cairoli [1], [2], [3] and Cairoli-Walsh [1].