

SÉMINAIRE DE PROBABILITÉS (STRASBOURG)

MARC YOR

PAUL-ANDRÉ MEYER

**Sur l'extension d'un théorème de Doob à un noyau
 σ -fini, d'après G. Mokobodzki**

Séminaire de probabilités (Strasbourg), tome 12 (1978), p. 482-488

<http://www.numdam.org/item?id=SPS_1978__12__482_0>

© Springer-Verlag, Berlin Heidelberg New York, 1978, tous droits réservés.

L'accès aux archives du séminaire de probabilités (Strasbourg) (<http://portail.mathdoc.fr/SemProba/>) implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques
<http://www.numdam.org/>

SUR L'EXTENSION D'UN THEOREME DE DOOB A UN NOYAU σ -FINI

par Marc YOR et P.A. MEYER, d'après G. MOKOBODZKI

1. Soient $(X, \underline{X})$ et $(Y, \underline{Y})$ deux espaces mesurables. La tribu $\underline{Y}$ est supposée séparable. On désigne par P une probabilité sur $(Y, \underline{Y})$, et par μ un noyau de X dans Y , de base P . Autrement dit, μ est une application $x \mapsto \mu^x$ de X dans l'ensemble des mesures positives sur $(Y, \underline{Y})$, telle que

(1) Pour tout $x \in X$, μ^x est absolument continue par rapport à P .

(2) Pour tout $A \in \underline{Y}$, l'application $x \mapsto \mu^x(A)$ est $\underline{X}$ -mesurable.

Nous supposons dans toute la suite que les mesures μ^x sont σ -finies sur $(Y, \underline{Y})$, et nous nous posons le problème de la validité du "théorème de Doob" :

Problème . Existe t'il une fonction positive $g(x, y)$, $\underline{X} \otimes \underline{Y}$ -mesurable, telle que l'on ait pour tout $x \in X$

$$\frac{d\mu^x}{dP} = g(x, \cdot) \text{ P-p.s. ?}$$

Il est bien connu que ce problème admet une réponse positive lorsque les mesures μ^x sont bornées : c'est là le théorème de Doob proprement dit, l'une des premières applications de la théorie des martingales. Rappelons rapidement comment on démontre cela. On représente $\underline{Y}$ comme $\bigvee_n \underline{Y}_n$, où $(\underline{Y}_n)$ est une suite de sous-tribus de $\underline{Y}$ engendrées par des partitions finies de plus en plus fines. On peut calculer explicitement

$$\left. \frac{d\mu^x}{dP} \right|_{\underline{Y}_n} = g_n(x, \cdot)$$

et vérifier que $g_n(\cdot, \cdot)$ est $\underline{X} \otimes \underline{Y}_n$ -mesurable. D'après le théorème de convergence des martingales, on peut alors prendre

$$g(x, y) = \liminf_{n \rightarrow \infty} g_n(x, y)$$

Mais ce raisonnement est entièrement en défaut lorsque les mesures μ^x ne sont pas bornées. En effet, il existe alors des exemples où l'on a, pour un certain x , $g_n(x, \cdot) = +\infty$ identiquement, et la fonction $g(x, \cdot)$ n'est donc pas une densité de la mesure σ -finie μ^x (un tel exemple figure dans [N], remarque suivant le corollaire II.2.12, au bas de la page 31). Il y a cependant un cas où l'extension est triviale : celui où le noyau μ est propre. Il existe alors une fonction $u(y)$ partout > 0 sur Y telle que $\mu^x(u) < \infty$ pour tout x , et l'on peut poser alors $\nu^x(dy) = u(y)\mu^x(dy)$,

choisir une version mesurable $h(x, \cdot)$ de la densité dv^x/dP , puis poser enfin

$$g(x, y) = \frac{1}{u(y)} h(x, y) .$$

Nous nous intéressons ici au cas général. Il semble y avoir une méthode d'attaque évidente. Puisque la mesure μ^x est σ -finie et absolument continue, on a

$$\mu^x = \lim_{n \rightarrow \infty} \mu^x \wedge (nP)$$

et on est amené au lemme suivant :

Lemme 1. Les conditions suivantes sont équivalentes

- a) Pour tout n , $x \mapsto \mu_n^x = \mu^x \wedge (nP)$ est un noyau.
- b) μ vérifie le théorème de Doob.

Démonstration. a) $\Rightarrow$ b). Le théorème de Doob étant vrai dans le cas borné, soit $g_n(x, y)$ une fonction $\underline{X} \otimes \underline{Y}$ -mesurable telle que $g_n(x, \cdot)$ soit densité de μ_n^x pour tout x . Alors $g(x, y) = \sup_n g_n(x, y)$ est mesurable, et $g(x, \cdot)$ est densité de μ^x pour tout x .

b) $\Rightarrow$ a). Soit $g(x, y)$ une fonction $\underline{X} \otimes \underline{Y}$ -mesurable telle que $g(x, \cdot)$ soit densité de μ^x pour tout x . Alors pour tout $A \in \underline{Y}$ on a

$$\mu_n^x(A) = \int_A g(x, y) \wedge n \, dP(y)$$

et il est évident que μ_n est un noyau.

2. On va maintenant regarder les choses de manière plus approfondie, et il nous faut pour cela quelques notations.

Nous désignons par $\overline{F}$ l'ensemble de toutes les classes de variables aléatoires positives, finies ou non, sur $(Y, \underline{Y})$, par F l'ensemble des classes de v.a. finies et positives, par F_b enfin l'ensemble des classes intégrables et positives ($F_b = L_+^1(\underline{Y}, P)$).

Il est bien connu que $\overline{F}$ est un espace polonais pour la topologie de la convergence en probabilité. Nous désignerons par Π la tribu borélienne correspondante sur $\overline{F}$, et aussi la tribu induite sur F ou F_b , le contexte se chargeant d'indiquer sur quel espace on se trouve.

Il y a une autre tribu intéressante sur $\overline{F}$, c'est la tribu engendrée par les applications $f \mapsto \int_A f dP$, où A parcourt $\underline{Y}$. Nous la désignerons par $\underline{T}$, et nous noterons de même par $\underline{T}$ les tribus induites sur F , F_b .

On a $\underline{T} \subset \Pi$. En effet, l'application $f \mapsto \int_A f dP = \sup_n \int_A f \wedge n \, dP$ est s.c.i., donc borélienne, pour la convergence en probabilité. Cela nous autorise à appeler $\underline{T}$ la tribu faible, Π la tribu forte, sur $\overline{F}$, F ou F_b .

Une application mesurable de $(X, \underline{X})$ dans $(F, \underline{T})$ est exactement ce que nous avons appelé un noyau (à mesures σ -finies, de base P) ; il nous arrivera de préciser en disant : noyau faible. Nous appellerons noyau fort une application mesurable de $(X, \underline{X})$ dans (F, Π) .

La proposition suivante nous indique quelle est la propriété de F_b qui "fait marcher" le théorème de Doob, et où se trouve le problème dans le cas général.

Proposition 1. a) Sur F_b , les tribus Π et T sont égales.

b) Un noyau μ satisfait au théorème de Doob si et seulement s'il est fort.

Démonstration. Il y a bien des manières de prouver a). L'argument suivant est rapide : on sait que sur un espace de Banach séparable, tel que $L^1(\underline{Y}, P)$, la tribu borélienne de la topologie de la norme coïncide avec celle de la topologie faible. Cette égalité s'induit sur $F_b = L^1_+$. Or sur F_b la tribu faible est $\underline{T}$, tandis que Π est intermédiaire entre les deux tribus précédentes.

Si μ est un noyau fort, $x \mapsto \mu^x \wedge (nP)$ est un noyau fort, car $f \mapsto f \wedge n$ est une application continue de F dans F . Donc μ satisfait au théorème de Doob d'après le lemme 1.

Inversement, supposons que $g(x, y)$ soit une fonction $\underline{X} \otimes \underline{Y}$ -mesurable positive, et que $\mu^x = g(x, \cdot)P$ pour tout x . Soit B la boule de centre $h \in F$ et de rayon a dans F

$$B = \{ f \in F : E[|f-h| \wedge 1] < a \}$$

Pour montrer que l'image réciproque de B par μ est $\underline{X}$ -mesurable, il suffit de vérifier que la fonction

$$x \mapsto \int_Y |g(x, y) - h(y)| \wedge 1 \, dP(y)$$

est mesurable, ce qui est évident. Donc μ est $(\underline{X}/\Pi)$ -mesurable.

3. Nous passons maintenant aux choses nouvelles, c'est à dire à l'existence de noyaux qui ne satisfont pas au théorème de Doob. Il suffit de démontrer pour cela que les tribus $\underline{T}$ et Π ne sont pas identiques en général, ce qui résulte du théorème suivant :

Proposition 2. Si l'espace probabilisé $(Y, \underline{Y}, P)$ est sans atomes, la tribu $\underline{T}$ n'est pas séparable. En particulier, on a $\underline{T} \neq \Pi$.

On déduit de là que $T = \Pi$ dans le seul cas où l'espace probabilisé est purement atomique.

Pour démontrer cette proposition, nous rappelons quelques résultats très élémentaires de théorie de la mesure. D'abord, l'absence d'atomes s'énonce ainsi :

(*) Pour tout $C \in \underline{Y}$ non négligeable, il existe $C' \in \underline{Y}$, $C' \subset C$ tel que $0 < P(C') < P(C)$.

On construit alors immédiatement une suite décroissante $C'_n \subset C$, avec $P(C'_{n+1}) < P(C'_n)$ pour tout n . Considérant les ensembles $C'_n \setminus C'_{n+1}$, on voit que :

(*) Pour tout $C \in \underline{Y}$ non négligeable et tout $\varepsilon > 0$, il existe $C_\varepsilon \in \underline{Y}$, $C_\varepsilon \subset C$ tel que $0 < P(C_\varepsilon) < \varepsilon$.

Au moyen de ces résultats, nous démontrons le lemme suivant (une démonstration très simple peut se déduire aussi de la note de Dellacherie sur le lemme de Borel-Cantelli, dans ce volume)

Lemme 2. Supposons $(Y, \underline{Y}, P)$ sans atomes. Soit $(C_n)_{n \geq 0}$ une suite d'éléments de $\underline{Y}$ non négligeables. Il existe alors $g \in F$ telle que

$$\int_{C_n} g \, dP = +\infty \text{ pour tout } n .$$

Démonstration. 1) Nous traitons le cas d'un seul $C \in \underline{Y}$ non négligeable. D'après (*) nous choisissons une suite décroissante (C'_k) d'éléments de $\underline{Y}$ tels que $C'_0 \subset C$, $P(C'_k) > P(C'_{k+1}) > 0$ pour tout k , et nous prenons g de la forme

$$g = \sum_k a_k 1_{C'_k \setminus C'_{k+1}}$$

où les a_k sont des constantes > 0 choisies de sorte que $\int_C g \, dP = +\infty$.

2) Donnons nous une fonction $u \in F$, un ensemble C comme ci-dessus, un nombre $\varepsilon > 0$, et montrons qu'il existe $v \in F$ telle que

$$\int_C v \, dP = +\infty, \quad v \geq u \text{ partout}, \quad P\{v \geq u + \varepsilon\} \leq \varepsilon .$$

A cet effet, nous choisissons d'après (*) un $C_\varepsilon \subset C$, $C_\varepsilon \in \underline{Y}$ tel que $0 < P(C_\varepsilon) < \varepsilon$, puis d'après 1) une fonction $g \in F$ telle que $\int_{C_\varepsilon} g \, dP = +\infty$. Nous prenons alors $v = u$ sur $C \setminus C_\varepsilon$, $v = g$ sur C_ε .

3) Démontrons enfin l'énoncé. Nous construisons par récurrence une suite $(g_n)_{n \geq 0}$ d'éléments de F tels que

$$\begin{aligned} - & g_0 = 1 \\ - & \int_{C_{n+1}} g_{n+1} \, dP = +\infty, \quad g_{n+1} \geq g_n, \quad P\{g_{n+1} \geq g_n + 2^{-n}\} \leq 2^{-n} \end{aligned}$$

ce qui est possible d'après 2). Nous posons $g = \lim_n g_n$; la dernière condition assure que $g \in F$, et g satisfait à l'énoncé.

Démonstration de la proposition 2. Supposons $\underline{T}$ séparable, et déduisons en une contradiction. $\underline{T}$ étant séparable et engendrée par les applications $f \mapsto \int_C f \, dP$, il existe une suite (C_n) d'éléments de $\underline{Y}$, que l'on

peut évidemment supposer non négligeables, tels que les applications $f \mapsto \int_{C_n} f dP$ engendrent $\underline{T}$. Alors la relation

$$(f \in F, \bar{f} \in F, \forall n \int_{C_n} f dP = \int_{C_n} \bar{f} dP) \text{ entraîne } (\forall A \in \underline{Y}, \int_A f dP = \int_A \bar{f} dP).$$

Soit g la fonction construite dans le lemme 2. Prenant $f=g$, $\bar{f}=g+1$, on voit que $\int_A g dP = \int_A (g+1) dP$ pour tout $A \in \underline{Y}$. Prenant $A=\{g \leq n\}$, on voit que $P(A)=0$, d'où $g=0$, ce qui est absurde.

4. Faisons une courte digression. Ce que la démonstration précédente montre, en fait, c'est que les applications $f \mapsto \int_A f dP$ ($A \in \underline{Y}$) séparent les points de F , mais qu'il n'existe aucune sous-famille dénombrable séparant les points de F . Il ne faut pas s'en étonner : la tribu Π est en effet une tribu de Blackwell sur F , comme on va le voir dans un instant. S'il existait une tribu séparable $\underline{T}' \subset \underline{T}$ séparant les points de F , on aurait $\underline{T}' = \underline{T} = \Pi$ d'après le théorème de Blackwell.

Lemme 3. F est borélien dans $\bar{F}$.

Démonstration. Pour toute $f \in F$ posons $k_n(f) = \int_Y \left(\frac{f}{1+f}\right)^n dP$. Les fonctions k_n sont continues sur F , et tendent en décroissant vers la fonction $k(f) = P\{f = +\infty\}$. L'ensemble $F = \{f : k(f) = 0\}$ est donc l'intersection des ensembles $\{f : k(f) < 1/m\}$, qui sont ouverts puisque $k = \inf_n k_n$ est s.c.s... Donc F est un G_δ dans $\bar{F}$, et en particulier F est polonais.

[Ce blanc est dû à C. Dellacherie, qui a trouvé la démonstration très simple ci-dessus, au lieu d'une démonstration compliquée de P.A. Meyer prouvant seulement que F était analytique dans $\bar{F}$].

5. Il semble difficile de reconnaître a priori si un noyau est un noyau fort. Aussi est-il intéressant d'étudier des variantes de cette condition, peut-être plus faciles à vérifier. Désignons par E l'image de $\underline{Y}$ par l'application $A \mapsto I_A$ de $\underline{Y}$ dans $L^\infty(\underline{Y}, P)$, et par $\underline{E}$ la tribu induite par L^∞ .

Définition. Un noyau μ de X dans Y (de base P , à mesures σ -finies), est dit mesurable si l'application $(x, A) \mapsto \mu^x(A)$ est $\underline{X} \otimes \underline{E}$ -mesurable.

Lemme 4. Tout noyau fort est mesurable.

Démonstration. Nous savons que, si μ est un noyau fort, il satisfait au théorème de Doob. Or soit $g(x,y)$ une fonction positive $\underline{X} \otimes \underline{Y}$ -mesurable. L'application

$$(x,A) \mapsto \int_A g(x,y) dP(y)$$

est $\underline{X} \otimes \underline{E}$ -mesurable. On se ramène en effet, par troncation, au cas où g est bornée ; puis, par classes monotones, au cas où $g(x,y) = u(x)v(y)$ ($u \in \text{ueb}(\underline{X})$, $v \in \text{veb}(\underline{Y})$), et le résultat est alors évident.

Inversement, un noyau mesurable est "presque" un noyau fort : il devient fort au prix d'un léger élargissement de la tribu $\underline{X}$.

Proposition 3. Soit μ un noyau mesurable de $(X, \underline{X})$ dans $(Y, \underline{Y})$, et soit $\hat{\underline{X}}$ la tribu engendrée par les ensembles $\underline{X}$ -analytiques. Alors μ est un noyau fort de $(X, \hat{\underline{X}})$ dans $(Y, \underline{Y})$.

Corollaire. Il existe une fonction $g(x,y)$ $\hat{\underline{X}} \otimes \underline{Y}$ -mesurable et positive, telle que $g(x, \cdot)$ soit une densité de $\mu^{\hat{\underline{X}}}$ pour tout $x \in X$.

Démonstration. Il nous suffit de démontrer que pour tout n , la fonction $x \mapsto \mu^{\hat{\underline{X}}} \wedge (nP)$ est un noyau de $(X, \hat{\underline{X}})$ dans $(Y, \underline{Y})$, ou encore, que pour tout $A \in \underline{Y}$, la fonction

$$x \mapsto nP(A) - \langle (nP - \mu^{\hat{\underline{X}}})^+, 1_A \rangle = nP(A) - \sup_{B \subset A} [nP(B) - \mu^{\hat{\underline{X}}}(B)]$$

est $\hat{\underline{X}}$ -mesurable. Or soit la fonction $\underline{X} \otimes \underline{E}$ -mesurable sur $\underline{X} \otimes \underline{E}$

$$j(x,B) = nP(B) - \mu^{\hat{\underline{X}}}(B)$$

et soit $h(x) = \sup_{B \subset A} j(x,B)$. L'ensemble $\{x \mid h(x) > a\}$ est la projection

sur X de l'ensemble

$$G = \{(x,B) : j(x,B) > a, B \subset A\}$$

Or l'ensemble des $B \in \underline{E}$ tels que $B \subset A$ est fermé dans $\underline{E}$, donc G est borélien, et sa projection sur X est analytique, et enfin $\hat{\underline{X}}$ -mesurable.

Remarque. La démonstration permet de prouver un peu mieux : si μ est un noyau mesurable, les fonctions $(x,A) \mapsto (\mu^{\hat{\underline{X}}} \wedge (nP))(A)$ sont $\underline{X} \otimes \underline{E}$ -coanalytiques au sens de [PP], III.61.

6. Nous avons vu plus haut qu'il existe des noyaux faibles qui ne sont pas forts. Le modèle de ces noyaux s'obtient en prenant $(X, \underline{X}) = (F, \underline{T})$, μ étant l'application identique de F sur F . Ce même modèle montre qu'il existe des noyaux faibles qui ne sont pas pas mesurables. En effet, que signifierait la mesurabilité de ce noyau ? Que l'application $(f,A) \mapsto \int f dP$ de $F \times \underline{E}$ dans $\underline{E}_+$ est $\underline{T} \otimes \underline{E}$ -mesurable. Or si elle l'était, elle le serait ^A

encore par rapport à la tribu sur $F \times E$ engendrée par une suite de rectangles $A_n \times B_n$ ($A_n \in \underline{T}$, $B_n \in \underline{E}$), et par conséquent par rapport à $\underline{S} \otimes \underline{E}$, où $\underline{S}$ est la tribu séparable engendrée par les A_n . Mais alors $\underline{T}$ serait égale à $\underline{S}$, et donc séparable, ce qui est faux (proposition 2).

REFERENCES

- [N] : J. Neveu , **Martingales à temps discret**, Masson, Paris, 1972.
 [PP] : C. Dellacherie et P.A. Meyer, Probabilités et Potentiel A, Hermann, Paris 1975.

L'article de G. Mokobodzki paraîtra sans doute au Séminaire d'Initiation à l'analyse (G. Choquet), ou au Séminaire de Théorie du Potentiel IV (Lecture Notes in M. Springer-Verlag).