Séminaire Brelot-Choquet-Deny. Théorie du potentiel

Paul-André Meyer

Théorèmes fondamentaux du calcul des probabilités (suite)

Séminaire Brelot-Choquet-Deny. Théorie du potentiel, tome 5 (1960-1961), exp. n° 2, p. 1-9

<http://www.numdam.org/item?id=SBCD_1960-1961__5__A3_0>

Séminaire BRELOT-CHOQUET-DENY
(Théorie du Potentiel)
5e année, 1960/61, n° 2

THÉORÈMES FONDAMENTAUX DU CALCUL DES PROBABILITÉS (suite)

par Paul-André MEYER

III. Probabilités et espérances mathématiques conditionnelles.

A. Définition des espérances conditionnelles.

THÉORÈME 3.1. - Soient p une loi de probabilité sur un espace mesurable (Ω, F), X une variable aléatoire numérique sur Ω, p-intégrable, et T une sous-tribu de F. Il existe une variable aléatoire Y, et une seule à une équivalence près, T-mesurable, p-intégrable, et telle que pour tout ensemble B de T, on ait :

$$\int_B Y(\omega)\, dp(\omega) = \int_B X(\omega)\, dp(\omega) \quad .$$

Une telle variable aléatoire Y est appelée espérance mathématique conditionnelle de X par rapport à T : on la note $E(X|T)$, ou parfois $E^T(X)$.

DÉMONSTRATION. - Commençons par supposer X de carré sommable : comme l'espace des fonctions T-mesurables de carré sommable est un sous-espace fermé de l'espace de Hilbert $L^2(\Omega, F, p)$, X admet une projection sur cet espace : soit Y cette projection, elle satisfait aux propriétés demandées.

Remarquons maintenant que si X est positive, Y est positive : il suffit de prendre pour B l'ensemble $\{Y < 0\}$ dans la formule (3.1.1) pour le voir. Il en résulte que si maintenant $X \in L^1(\Omega, F, p)$, et si $X \geq 0$ presque sûrement (p. s.), les projections Y_n des variables aléatoires $X_n = \inf(X, n)$ croissent p. s. : elles convergent donc vers une variable aléatoire Y qui satisfait à (3.1.1). Il ne reste plus qu'à appliquer ce résultat à X^+ et X^-, si X est quelconque, pour obtenir le théorème.

Reste l'unicité de Y. Elle est triviale, car si Y_1 et Y_2 sont deux solutions, en prenant pour B dans (1) les ensembles $\{Y_1 - Y_2 > 0\}$, $\{Y_1 - Y_2 < 0\}$, on obtient l'égalité p. s. de Y_1 et Y_2.

B. Propriétés fondamentales des espérances conditionnelles.

1° Si X est une constante, $E^T(X) = X$ p. s. L'application $X \to E^T(X)$ est linéaire.

2° Si X est positive, $E^T(X) \geq 0$ p. s.

3° Si X est T-mesurable, $E^T(X) = X$ p. s.

4° $E[E^T(X)] = E(X)$.

5° Supposons que Y soit T-mesurable, et par exemple bornée : on a $\forall X$:

$$E^T(XY) = Y.E^T(X) \text{ p. s.}$$

(Commencer par supposer X et Y positives, le prouver pour Y étagée, puis passer à la limite en utilisant une suite croissante de Y_n étagées).

6° Si U est une sous-tribu de T , on a :

$$E^U E^T(X) = E^U(X) \text{ p. s.}$$

NOTATIONS. - Si T est la tribu engendrée par une famille de fonctions $\{f_i\}_{i\in I}$, nous écrirons : $E(X|f_i , i \in I)$.

Si X est la fonction caractéristique d'un ensemble A , la fonction $E(X|T)$ prend le nom de probabilité conditionnelle de l'événement A par rapport à T , et se note $P(A|T)$ ou $P^T(A)$: il importe de se rappeler qu'une probabilité conditionnelle est une classe de variables aléatoires équivalentes, et non pas un nombre !

C. Définition de l'indépendance conditionnelle.

Soit (Ω , F , p) un espace muni d'une loi de probabilité, et soient B_i , $i = 1 , 2 , 3$, trois sous-tribus de F : nous dirons que B_1 et B_3 sont conditionnellement indépendantes relativement à B_2 (si B_2 était la tribu engendrée par une application f , on dirait : f étant donnée) si, pour tout couple de variables aléatoires Y_i ($i = 1 , 3$) B_i-mesurables, et de carré sommable, on a la relation

$$E(Y_1 Y_3|B_2) = E(Y_1|B_2).E(Y_3|B_2) \quad . \tag{1}$$

Il suffit évidemment que cette égalité ait lieu lorsque Y_1 et Y_3 sont des fonctions caractéristiques d'ensembles.

THÉORÈME 3.2. - Pour que B_1 et B_3 soient conditionnellement indépendantes relativement à B_2 , il faut et il suffit que, B_{12} étant la tribu engendrée par B_1 et B_2 , on ait :

$$E(Y_3|B_{12}) = E(Y_3|B_2) \tag{2}$$

pour toute variable aléatoire intégrable Y_3 , B_3-mesurable.

DÉMONSTRATION.

(2) $\Rightarrow$ (1) : $E^{B_2}(Y_1\ Y_3) = E^{B_2}(E^{B_{12}}(Y_1\ Y_3))$ puisque $B_2 \subset B_{12}$ (propriété 6°) = $E^{B_2}(Y_1\ E^{B_{12}}(Y_3))$ (Y_1 est B_{12}-mesurable propriété 5°) = $E^{B_2}(Y_1\ E^{B_2}(Y_3))$ d'après la relation (2) = $E^{B_2}(Y_1)\ E^{B_2}(Y_3)$ (propriété 5°).

C. Q. F. D.

(1) $\Longrightarrow$ (2) : il suffit de vérifier que les deux membres de (2) ont même intégrale sur tout ensemble de la forme $A_1 \cap A_2$ de B_{12}, où $A_1 \in B_1$, $A_2 \in B_2$. Soient a_1, a_2 leurs fonctions caractéristiques.

$E(a_1\ a_2\ E^{B_2}(Y_3)) = E(E^{B_2}(a_1\ a_2\ E^{B_2}(Y_3)))$ (propriété 4°) $= E(a_2\ E^{B_2}(a_1\ E^{B_2}(Y_3)))$ (propriété 5°) $= E(a_2\ E^{B_2}(Y_3)\ E^{B_2}(a_1))$ (propriété 5°) $= E(a_2\ E^{B_2}(a_1\ Y_3))$ d'après la relation (1) $= E(E^{B_2}(a_1\ a_2\ Y_3))$ (propriété 5°) $= E(a_1\ a_2\ Y_3)$ (propriété 4°)

C. Q. F. D.

D. REMARQUE.- Soient (Ω , F , p) un espace muni d'une loi de probabilité, X une variable aléatoire numérique intégrable sur cet espace, (E , T) un espace mesurable, $f : (\Omega , F) \to (E , T)$ une application mesurable, et q la mesure image de p par f : on démontre très facilement, comme dans le théorème 3.1, qu'il existe une variable aléatoire Y sur (E , T) et une seule à une équivalence près, telle que, pour toute partie B de T :

$$\int_B Y\, dq = \int_{f^{-1}(B)} X\, dp$$

il est clair que $Y \circ f$ est une version de $E(X|f)$: la fonction $y \to Y(y)$ sur E s'appelle l'espérance mathématique conditionnelle de X, sachant que $f(x) = y$. Cette terminologie est parfois très commode.

IV. Définition, propriétés générales, construction des processus stochastiques.

A. Donnons-nous un ensemble d'indices $\mathcal{T}$: ce sera ici une partie de $\overline{R}$, généralement un intervalle de R ou de Z. Le point t de $\mathcal{T}$ sera appelé "l'instant t". Soit (E , B) un espace mesurable, qui sera appelé espace des états : on appelle processus stochastique (à valeurs dans (E , B)) un système formé :

- d'un espace (Ω , F , p) muni d'une loi de probabilité, appelé espace de base.

- d'une famille d'applications mesurables $\{X_t\}_{t\in\mathcal{T}}$ définies sur l'espace de base, à valeurs dans l'espace des états.

Soit $\omega \in \Omega$: l'application $t \to X_t(\omega)$ est appelée la trajectoire associée à ω.

DÉFINITION 4.1. - Soient $(\Omega , F , p , \{X_t\}_{t\in\mathcal{T}})$, $(\Omega' , F' , p' , \{X'_t\}_{t\in\mathcal{T}})$ deux processus stochastiques ayant même espace d'états, et même ensemble d'indices. On dit que ces deux processus sont équivalents si, pour tout système fini $u = (t_1 , t_2 \dots t_n)$ d'instants de t , les répartitions des variables aléatoires $(X_{t_1} \dots X_{t_n})$, $(X'_{t_1} \dots X'_{t_n})$ dans l'espace produit (E^u , B^u) sont identiques.

Soit U la famille des parties finies de $\mathcal{T}$: si $u \in U$, nous noterons N_u la projection de $E^{\mathcal{T}}$ sur E^u , et si $u \subset v \in U$, N_{uv} la projection de E^v sur E^u .

DÉFINITION 4.2. - On appelle système projectif de lois de probabilités sur $(E^{\mathcal{T}} , B^{\mathcal{T}})$ une famille de lois de probabilités π_u sur les espaces $(E^u , B^u)_{u\in U}$, telles que, si $u \subset v \in U$, l'on ait $N_{uv}(\pi_v) = \pi_u$. S'il existe une loi de probabilité π sur $(E^{\mathcal{T}} , B^{\mathcal{T}})$ telle que $\forall$ u , $N_u(\pi) = \pi_u$, on dit que π est la limite projective des lois π_u .

Il est immédiat qu'il existe au plus une limite projective. Toute classe d'équivalence de processus stochastiques détermine un système projectif de lois de probabilité, qui admet une limite projective : π_u étant une loi image de la loi du processus par l'application $\omega \to \{X_t(\omega)\}_{t\in u}$ de (Ω , F) dans (E^u , B^u) , et la limite projective π des π_u étant l'image de la loi du processus par l'application $\omega \to \{X_t(\omega)\}_{t\in\mathcal{T}}$ de (Ω , F) dans $(E^{\mathcal{T}} , B^{\mathcal{T}})$. Le processus ainsi construit, $(E^{\mathcal{T}} , B^{\mathcal{T}} , \pi , \{Y_t\}_{t\in\mathcal{T}})$ - où les Y_t désignent maintenant les applications coordonnées - sera appelé premier processus canonique associé à la classe d'équivalence donnée.

Le problème inverse (étant donné un système projectif de lois de probabilité, est-il le système des répartitions finies d'un processus stochastique ?) est évidemment très important, et équivaut au problème de l'existence d'une limite projective pour un tel système. On a des exemples qui prouvent qu'il n'existe pas toujours de telle limite projective lorsque les mesures sont abstraites. En revanche, on a le théorème suivant :

THÉORÈME 4.1. - Si E est un espace compact, et si les π_u sont des lois de Radon, il existe une loi de Radon π sur l'espace compact $E^{\mathcal{T}}$ telle que $N_u(\pi) = \pi_u$. (cf. [1], p. 100).

En général, notre espace E ne sera pas compact, mais, à cause de la nature particulière des processus que nous étudierons, il n'y aura pas d'inconvénient à le compactifier de manière à pouvoir appliquer ce théorème. La tribu $M(\pi)$ des ensembles mesurables pour la mesure de Radon π contenant évidemment la tribu $B^{\mathcal{T}}$, l'existence d'une limite projective sur $B^{\mathcal{T}}$ en résulte. Nous appelerons <u>second processus canonique</u> le processus $(E^{\mathcal{T}}, M(\pi), \pi, \{X_t\}_{t\in T})$.

Nous utiliserons désormais des notations simplifiées : nous omettrons la mention de l'espace Ω, de F, de $\mathcal{T}$ ... lorsque cela n'introduira pas d'ambiguité : nous parlerons par exemple du "processus $\{X_t\}$" ...

B. <u>Processus séparables</u>. - Soient $\{X_t\}$ et $\{Y_t\}$ deux processus équivalents : Si $\mathcal{T}$ n'est pas dénombrable, il se peut que, pour les lois correspondantes, des "événements" au sens intuitif du terme (par exemple "l'événement" $\{\forall\, t \in \mathcal{T}, X_t(\omega) \in A\}$, où A est une partie mesurable de E), ne soient pas mesurables, ou bien le soient, et se voient attribuer des mesures différentes pour les deux processus. Il convient donc de donner un nom à certains processus pour lesquels certains "événements" intéressants dont la définition fait intervenir une infinité plus que dénombrable d'instants, ont des probabilités naturelles.

Nous supposerons dans les théorèmes qui suivent que $\mathcal{T} = R_+$: l'adaption à d'autres cas serait triviale.

K étant un compact de E, I une partie quelconque de R_+ nous noterons $V(I, K)$ l'ensemble des ω tels que $\forall\ t \in I$ on ait $X_t(\omega) \in K$.

DÉFINITION 4.3. - On dit qu'un processus $(\Omega, F, p, \{X_t\})$ à valeurs dans un espace LCD E est <u>séparable</u> (relativement à la famille des parties compactes de E) si, pour toute partie compacte K de E, et tout intervalle ouvert I de R_+, l'ensemble $V(I, K)$ appartient à F, et si l'on a :

$$p[V(I, K)] = \inf_{U} p[V(U, K)] \quad \text{où } U \text{ parcourt la famille des parties finies de } I$$

Il est évident qu'il existe alors une partie dénombrable S de I telle que $p[V(S, K)]$ soit égal à l'inf ci-dessus : une telle partie sera appelée <u>ensemble séparant</u> (relatif à I et K) : s'il existe une partie dénombrable S qui est séparante relativement à tout couple (I , K) , nous dirons que S est un ensemble séparant universel.

LEMME. - Pour tout processus stochastique, à valeurs dans un espace LCD, il existe un ensemble séparant universel.

Prenons en effet une suite de compacts K_n telle que tout compact soit l'intersection d'une sous-suite des K_n (par exemple : les réunions finies de boules fermées dont le rayon est rationnel, et dont le centre appartient à un ensemble dénombrable partout dense dans X) . Prenons une suite d'intervalles ouverts I_m , formant une base de la topologie de R_+ . Pour chaque couple (m , n), soit S_{mn} un ensemble séparant relativement à I_m , K_n : $S = \cup S_{mn}$ répond à la question. Nous supposerons dans la suite que S est en outre dense dans R_+ .

THÉORÈME 4.2. - Pour tout processus à valeurs dans un espace compact E , il existe un processus équivalent qui est séparable.

DÉMONSTRATION. - Montrons que le second processus canonique associé au processus donné est séparable : soit K une partie compacte de E : les ensembles $V(I , K)$, où I est quelconque, sont compacts dans E^{R_+} . Ils appartiennent donc à la tribu $M(p)$ définie plus haut. Les ensembles $V(U , K)$, où U parcourt la famille des parties finies de I , forment une famille filtrante décroissante de compacts dont l'intersection est $V(I , K)$, et il en résulte que

$$p[V(I , K)] = \inf_U p[V(U , K)] \quad .$$

C. Propriétés des processus séparables.

Faisons une remarque qui nous sera utile dans la suite ; tout processus dont les trajectoires sont p. s. continues à droite est séparable, et tout ensemble dense est un ensemble séparant universel pour un tel processus.

Une définition équivalente de la séparabilité, souvent précieuse, est la suivante :

THÉORÈME 4.3. - Pour toute partie A de $\mathcal{T}$, tout $\omega \in \Omega$, posons

$$X_A(\omega) = \{X_t(\omega) , t \in A\} \quad .$$

Pour que le processus $\{X_t\}$ à valeurs dans l'espace métrique compact E , soit séparable, il faut et il suffit que pour presque tout ω , on ait :

$$\overline{X_I(\omega)} = \overline{X_{I\cap S}(\omega)}$$

pour tout intervalle ouvert I, S désignant un ensemble séparant universel.

DÉMONSTRATION. - Il est clair que ceci entraîne bien la séparabilité. Inversement, montrons que, si le processus est séparable, cette propriété a lieu pour les intervalles I_m à extrémités rationnelles ; on en déduira immédiatement qu'elle est vraie pour des intervalles ouverts quelconques. Soit K_n une suite de compacts telle que tout compact K soit l'intersection d'une sous-suite des K_n : soit W_{nm} l'ensemble de mesure nulle $V(I_m \cap S, K_n) - V(I_m, K_n)$: si $W = \bigcup W_{nm}$, la condition de l'énoncé est remplie pour tout $\omega \notin W$. On en déduit :

THÉORÈME 4.4. - Soit X_t un processus séparable à valeurs dans l'espace métrique compact E, et soit S un ensemble séparant universel : pour que les trajectoires soient p. s. continues à droite à l'instant t, il faut et il suffit que $X_s(\omega) \underset{\substack{s \to t_+ \\ s \in S}}{\to} X_t(\omega)$ p. s. L'ensemble des trajectoires qui possèdent en tout point t de R_+ une limite à droite et une limite à gauche est mesurable, et identique à l'ensemble des trajectoires ω telles que, pour tout $t \in \mathcal{T}$, $\lim\limits_{\substack{s \to t_+ \\ s \in S}} X_s(\omega)$ et $\lim\limits_{\substack{s \to t_- \\ s \in S}} X_s(\omega)$ existent.

DÉMONSTRATION. - Nous supposons ici que $E = \overline{R}$ (Cf. Appendice). L'existence d'une limite à droite au point t s'exprime ainsi : la fonction $s \to X_s(\omega)$ possède une valeur d'adhérence à droite au point t. On applique alors le théorème 4.3.

Soit I un intervalle ouvert relativement compact, U une partie finie de I, et soient r_1, r_2 deux nombres rationnels, $r_1 < r_2$. Appelons nombre des passages décroissants de la trajectoire associée à ω sur l'intervalle $[r_1, r_2]$, aux instants de U, le nombre $n(\omega, U, r_1, r_2)$ des couples d'instants u_i, u_j de u tels que : $u_i < u_j$; $X_{u_i}(\omega) > r_2$; si $u_i < u < u_j$, $u \in U$, $r_1 \leq X_u(\omega) \leq r_2$; $X_{u_j}(\omega) < r_1$. Si maintenant V est une partie quelconque de I, posons

$n(\omega, V, r_1, r_2) = \sup\limits_{U \subset V} n(\omega, U, r_1, r_2)$. Or il est facile de voir, d'après le théorème 4.3, que $\forall\ U \subset I$, $n(\omega, U, r_1, r_2) \leq n(\omega, S \cap I, r_1, r_2)$ p. s., de sorte que $n(\omega, I, r_1, r_2) = n(\omega, S \cap I, r_1, r_2)$ p. s. Le premier membre est donc une variable aléatoire. L'absence de discontinuités de seconde espèce s'exprimant par la condition $n(\omega, I, r_1, r_2) < \infty$ p. s. $\forall\ I, r_1, r_2$, le théorème est démontré.

THÉORÈME 4.5. - Soit (Ω, F, p, X_t), où $t \in R_+$, un processus stochastique à valeurs dans un espace LCD, et S un ensemble dénombrable dense dans R_+. Supposons que :

1° $\forall\ t \in R_+$, $\lim\limits_{\substack{s \in S \\ s \to t_+}} X_s = X_t$ p. s.

2° L'ensemble des trajectoires dont la restriction à S admet une limite à droite et une limite à gauche en tout point de R_+ a pour mesure 1 . Il existe alors un processus équivalent, dont toutes les trajectoires sont en tout point de R_+ continues à droite et pourvues de limites à gauche.

DÉMONSTRATION. - Nous avons vu dans la démonstration précédente que l'ensemble des applications de S dans R_+ qui sont pourvues en tout point de R_+ de limites à droite et à gauche est mesurable sur la tribu B^S de E^S . Notre hypothèse est qu'ici il a la mesure 1 , pour le premier processus canonique associé au processus $\{X_s\}_{s \in S}$: choisissons-le comme espace de base, et posons, si $t \in R_+$:

$$\hat{X}_t(\omega) = \lim_{\substack{s \to t_+ \\ s \in S}} \hat{X}_s(\omega),$$ où les $\hat{X}_s$ sont les restrictions à l'ensemble des coordonnées de E^S .

Le processus $\hat{X}_t$ est équivalent au processus X_t , d'après la condition 1 .

BIBLIOGRAPHIE

[1] BOURBAKI (Nicolas). - Intégration, chap. 1-4. - Paris, Hermann, 1952 (Act. scient. et ind., 1175 ; Eléments de Mathématique , 13).

[2] DOOB (J. L.). - Stochastic processes. - New York, J. Wildy, 1953.

APPENDICE.

1. Corollaire du théorème 4.3.

Soient $(\Omega , F , p , \{X_t\}_{t\in\mathcal{T}})$ un processus séparable à valeurs dans un espace compact métrisable E , et S un ensemble séparant universel. Soit φ une application continue de E dans un espace compact métrisable G . Le processus $\{\varphi \circ X_t\}_{t\in\mathcal{T}}$ est séparable, et admet S pour ensemble séparant universel.

DÉMONSTRATION. - C'est une conséquence immédiate du théorème 4.3.

2. Fin de la démonstration du théorème 4.4.

Cette démonstration, telle qu'elle est donnée, n'a de sens que si $E = \overline{R}$. Voici comment on peut en déduire le cas où E est un compact métrisable : soit g_n une suite de fonctions numériques continues sur E , qui sépare les points de E . On peut appliquer le théorème 4.4 aux processus, à valeurs dans $\overline{R}$, $\{g_n \circ X_t\}_{t\in\mathcal{T}}$, qui sont séparables d'après le corollaire ci-dessus. Il en résulte bien que le théorème 4.4 est vrai pour le processus $\{X_t\}$.