ANNALES DE L'I. H. P., SECTION B

BERNARD BRU
HENRI HEINICH

Meilleures approximations et médianes conditionnelles

Annales de l'I. H. P., section B, tome 21, nº 3 (1985), p. 197-224
<http://www.numdam.org/item?id=AIHPB_1985__21_3_197_0>

Ann. Inst. Henri Poincaré,
Vol. 21, n° 3, 1985, p. 197-224. *Probabilités et Statistiques*

Meilleures approximations et médianes conditionnelles

par

Bernard BRU et Henri HEINICH
Université Paris VI[e], Laboratoire de probabilités,
4, Place Jussieu, 75230 Paris Cedex 05

RÉSUMÉ. — Nous étudions l'existence et les propriétés des meilleurs approximants d'une variable X d'un espace de Köthe général par rapport à une sous-tribu. En particulier si $\mathbb{E}$ est un Banach réticulé faiblement séquentiellement complet et $\mathbb{F}$ un sous-Banach réticulé, nous montrons pour tout point de $\mathbb{E}$ l'existence d'un meilleur approximant dans $\mathbb{F}$.

Dans une seconde partie, nous étudions la question des minima des fonctions $Y \to E\Phi(X - Y)$, pour une fonction Φ convexe nulle en zéro, absolument générale. Nous en déduisons le comportement des suites de meilleurs approximants d'une v. a. X le long d'une filtration : théorèmes de convergence, inégalités maximales, etc.

ABSTRACT. — We study existence and properties of best approximants of a random variable X in a general Köthe space with respect to a given sub σ-field. In particular if $\mathbb{E}$ is a weakly sequentially complete Banach lattice and $\mathbb{F}$ a sub-Banach lattice, we show, for every point of $\mathbb{E}$, the existence of a best approximant in $\mathbb{F}$.

In a second part, we study minima of functions $Y \to E\Phi(X - Y)$ for a general, null at zero, convex function Φ. We then give some properties of sequence of best approximants of a r. v. X following a given filtration.

Key-words: STMA: 00150-01010-01110
AMS: 41 A 65-60 G 48.

INTRODUCTION

Approcher de la meilleure façon possible, pour un critère global donné, une variable aléatoire inconnue X par une fonction $\psi(X_1, \ldots, X_n)$ de n variables aléatoires connues $X_1, \ldots, X_n$ n'est pas un problème très nouveau.

En 1774, dans son mémoire sur la probabilité des causes par les événements, Laplace avait déjà posé et résolu un tel problème en adoptant le critère que

$$\| X - \psi(X_1, \ldots, X_n) \|_1 \qquad \text{soit un minimum},$$

critère qui lui semblait « tenir à la nature du problème ». Il faisait toutefois observer qu'il était possible d' « imposer une infinité d'autres conditions semblables, qui donneront chacune un milieu différent » ([*19*] p. 230 ou [*20*] p. 477).

Laplace montrait que la meilleure approximation de X était dans ce cas une médiane de la loi conditionnelle de X sachant $(X_1, \ldots, X_n)$ (la loi *a posteriori* dans le cadre bayésien où il se plaçait).

Il est curieux de noter que le résultat de Laplace a été présenté dans sa version moderne exactement deux cents ans plus tard par Ando et Shintani dans leur article [*2*].

Entre temps, sur la proposition de Gauss en 1821, les savants avaient adopté un critère beaucoup plus maniable qui consistait à minimiser

$$\| X - \psi(X_1, \ldots, X_n) \|_2 .$$

L'espérance conditionnelle de X sachant $(X_1, \ldots, X_n)$ qui résoud ce problème, possède, en effet, entre autres avantages, la propriété d'être linéaire en X (ce que la médiane n'est pas).

Bienaymé avait d'ailleurs montré, en 1853, que tout autre critère, par exemple minimiser les normes 3 ou 4, conduisait à des complications inutiles.

Dans les années 60 (du 20ème siècle), pour diverses raisons théoriques ou pratiques, certains auteurs reconsidérèrent le problème général de la meilleure approximation d'une variable X par une variable Y mesurable par rapport à une sous-tribu donnée, pour une norme ou une fonctionnelle différentes de la norme 2. Citons l'article d'Amemiya et Ando [*1*], qui étudient le cas des normes L^p, $1 < p < \infty$ et les travaux de M. M. Rao sur les meilleures approximations dans les espaces d'Orlicz L^Φ, [*10*] [*24*] [*25*].

Ces auteurs ont notamment montré que les théorèmes classiques de la théorie des martingales s'étendent, dans une certaine mesure, aux suites de meilleurs approximants $(p_{\mathscr{B}_n}(X))$ d'une variable X par des variables $\mathscr{B}_n$-mesurables, pour une filtration $(\mathscr{B}_n)$ donnée.

Tout récemment Landers, Rogge et Herrndorf, dans une série d'articles [*13*] [*15*] [*16*] [*17*], ont précisé et étendu les travaux précédents [*12*] en traitant complètement le cas des fonctionnelles $E\Phi(|X - Y|)$ et des normes de Luxemburg des espaces d'Orlicz, dans le contexte très général des σ-lattices.

Nous nous proposons ici de poursuivre ce type d'étude dans deux directions différentes.

Nous commençons par étudier les meilleurs approximants d'une variable X par des variables Y mesurables par rapport à une sous-tribu pour une norme latticielle aussi générale que possible. Nous montrons, en utilisant les propriétés des espaces de Köthe, que de tels approximants existent sous des hypothèses très générales englobant tous les résultats antérieurs (Théorème 7).

Dans une seconde partie nous étudions la question particulière des minima des fonctions du type $Y \to E\Phi(X - Y)$, Y mesurable par rapport à une sous-tribu, Φ convexe. Nous montrons, en exploitant l'analogie qui existe entre ce problème et celui des M-estimateurs de la statistique, que ces minima sont toujours des sortes de médianes conditionnelles. Cette représentation nous permet d'obtenir la forme générale des inégalités maximales pour les suites et des inégalités de Doob de meilleurs approximants (proposition 17, 18, 19). Ces résultats étendent ceux obtenus par Landers et Rogge dans le cas des espaces L^p et les critères développés par les auteurs dans [*7*].

* * *

Nous remercions très vivement le rapporteur pour ses nombreuses et judicieuses observations, dont nous avons essayé de tenir le plus grand compte.

I. MEILLEURES APPROXIMATIONS DANS LES ESPACES DE KÖTHE

Soit $\mathbb{E}$ un espace de Banach de variables aléatoires définies sur un espace de probabilité $(\Omega, \mathfrak{A}, P)$. Nous dirons que $\mathbb{E}$ est un *espace de Köthe* si

i) On a $L^\infty \subset \mathbb{E} \subset L^1$ les injections étant continues.

ii) X mesurable, $Y \in \mathbb{E}$ et $|X| \leqq |Y|$ impliquent $X \in \mathbb{E}$ et $\|X\|_{\mathbb{E}} \leqq \|Y\|_{\mathbb{E}}$.

Nous adopterons pour ce type d'espace la terminologie et les notations de [*21*]. En particulier nous notons $\mathbb{E}^*$ le dual topologique de $\mathbb{E}$ et $\mathbb{E}' = \{ Y \in L^1 ; YX \in L^1, \forall X \in \mathbb{E} \}$. Nous dirons que $\mathbb{E}$ a la *propriété de Fatou,* si $X_n \uparrow X$ p. s., $X_n \in \mathbb{E}$ et $\| X_n \|_{\mathbb{E}} \leq 1$ impliquent $X \in \mathbb{E}$ et $\lim \| X_n \|_{\mathbb{E}} = \| X \|_{\mathbb{E}}$. Cette propriété est équivalente à $(\mathbb{E}')' = \mathbb{E}$ ([*21*]-I-*b*-18). $\mathbb{E}$ est à *norme continue pour l'ordre,* en abrégé n. c. o., si $X_\alpha \in \mathbb{E}_+$ et $X_\alpha \downarrow 0$ impliquent $\| X_\alpha \|_{\mathbb{E}} \downarrow 0$. Un espace de Köthe est n. c. o. si et seulement si $\mathbb{E}^* = \mathbb{E}'$ ([*21*]-I-*b*-17).

Tous les espaces L^p pour $1 \leqslant p < +\infty$, sont évidemment de ce type; L^∞, qui a la propriété de Fatou, n'est pas n. c. o. Les espaces d'Orlicz L^Φ ont la propriété de Fatou; ils sont n. c. o. si et seulement si Φ est modérée (Δ_2) c'est-à-dire si $\operatorname*{Sup}_{x>0} \dfrac{\Phi(2x)}{\Phi(x)} < +\infty$.

Soit $\mathbb{E}$ un espace de Köthe de v. a. r. sur $(\Omega, \mathfrak{A}, P)$ et soit $\mathscr{B}$ une sous-tribu de $\mathfrak{A}$. Nous notons $\mathbb{E}(\mathscr{B})$ l'ensemble des v. a. r. $\mathscr{B}$-mesurables de $\mathbb{E}$: $\mathbb{E}(\mathscr{B})$ est à son tour un espace de Köthe sur $(\Omega, \mathscr{B}, P)$. Pour $X \in \mathbb{E}$ nous dirons que $Z \in \mathbb{E}(\mathscr{B})$ est un *meilleur approximant de* X *sachant* $\mathscr{B}$ si on a : $\| X - Z \|_{\mathbb{E}} \leqslant \| X - Y \|_{\mathbb{E}}$ pour toute v. a. $Y \in \mathbb{E}(\mathscr{B})$.

Nous nous proposons d'étudier l'existence et les propriétés de tels meilleurs approximants. Au préalable établissons quelques résultats techniques.

1. Faible compacité dans les espaces de Köthe.

LEMME 1. — *Soit* $\mathbb{E}$ *un espace de Köthe ayant la propriété de Fatou et soit* (X_n) *une suite bornée en norme de v. a. de* $\mathbb{E}$, *convergeant faiblement dans* L^1 *vers une v. a.* X. *Alors on a* $X \in \mathbb{E}$ *et* $\| X \|_{\mathbb{E}} \leqslant \varliminf \| X_n \|_{\mathbb{E}}$.

Démonstration. — Si, pour la convergence faible de L^1, $X_n \to X$ et $| X_n | \to X'$ il est clair que $| X | \leqslant X'$. Il suffit donc de considérer des suites positives. Soit $Y \in \mathbb{E}'_+$, il s'agit d'abord de montrer que $XY \in \mathbb{L}^1$. Il existe une suite croissante $Y_p \in L^\infty_+$ convergeant p. s. vers Y. On a ainsi $\| Y_p \|_{\mathbb{E}^*} \leqslant \| Y \|_{\mathbb{E}^*}$ et $Y_p \to Y$ pour la topologie $\sigma(\mathbb{E}^*, \mathbb{E})$. Comme

$$E(X_n Y_p) \leqslant \operatorname*{Sup}_n \| X_n \|_{\mathbb{E}} \, . \, \| Y \|_{\mathbb{E}^*} \, .$$

On a $XY \in L^1$. Enfin $E(XY) \leqslant \varepsilon + E(XY_p) \leqslant 2\varepsilon + E(X_n Y_p)$ pour p et n assez grands, ce qui montre que $E(XY) \leqslant \varliminf E(X_n Y)$, pour tout $Y \in \mathbb{E}'$. Pour finir, il reste à observer que, $\mathbb{E}$ ayant la propriété de Fatou, $\mathbb{E}'$ est normant dans $\mathbb{E}$ ([*21*]-I-*b*-18) c'est-à-dire

$$\| X \|_{\mathbb{E}} = \operatorname{Sup} \{ E(XY), \quad Y \in \mathbb{E}', \quad \| Y \|_{\mathbb{E}^*} \leqslant 1 \} \leqslant \varliminf \| X_n \|_{\mathbb{E}} \, .$$

Remarque. — Si dans les hypothèses du lemme 1 nous remplaçons la convergence faible de X_n par une convergence p. s. les conclusions demeurent. En effet soit (X_n) une suite de v. a. positives de $\mathbb{E}$, bornées en norme et convergeant p. s. vers X et soit $Z \in L^\infty_+$. Le lemme de Fatou montre qu'on a $E(ZX) \leqslant \underline{\lim}\, E(ZX_n)$. Si Z_p est une suite croissante de L^∞_+ convergeant p. s. vers une v. a. positive Y appartenant à $\mathbb{E}'_+$; on peut écrire, à ε fixé et pour p et n grands

$$E(XY) \leqslant \varepsilon + E(XZ_p) \leqslant 2\varepsilon + E(X_n Z_p) \leqslant 2\varepsilon + \operatorname*{Sup}_n \| X_n \|_{\mathbb{E}} \,.\, \| Y \|_{\mathbb{E}^*} .$$

D'où $XY \in L^1$ et, comme précédemment $\| X \|_{\mathbb{E}} \leqslant \underline{\lim}\, \| X_n \|_{\mathbb{E}}$.

LEMME 2. — *Si $\mathbb{E}$ est un espace de Köthe on a équivalence entre :*

i) $\mathbb{E}$ est faiblement séquentiellement complet.

ii) $\mathbb{E}$ est n. c. o. et possède la propriété de Fatou.

La démonstration de ce lemme est classique ([*21*]-I-*c*-4).

Le résultat suivant est une conséquence aisée des lemmes précédents.

COROLLAIRE 3. — *Soit $\mathbb{E}$ un espace de Köthe faiblement séquentiellement complet tel que L^∞ soit fortement dense dans $\mathbb{E}^*$. Une partie bornée de $\mathbb{E}$ est faiblement relativement compacte si et seulement si elle l'est dans L^1.*

On a aussi

COROLLAIRE 4. — *Pour un espace de Köthe $\mathbb{E}$ faiblement séquentiellement complet, les propriétés suivantes sont équivalentes :*

i) (X_n) est une suite de v. a. faiblement relativement compacte de $\mathbb{E}$.

ii) $(|X_n|)$ est faiblement relativement compacte dans $\mathbb{E}$.

iii) (X_n^+) et (X_n^-) sont faiblement relativement compactes dans $\mathbb{E}$.

Démonstration i) $\Rightarrow$ ii). — Soit $Y \in \mathbb{E}^*_+$. Sous l'hypothèse (*i*) la suite $(X_n Y)$ étant équi-intégrable, il en est de même pour $(|X_n|Y)$. La suite $(|X_n|)$ étant elle-même équi-intégrable, ses points d'accumulation pour $\sigma(L^1, L^\infty)$ sont dans $\mathbb{E}$ — lemme 1 — et si $|X_{n_i}| \to Z$ $\sigma(L^1, L^\infty)$ on a $|X_{n_i}|Y \to ZY$, pour la même topologie. De là pour $\sigma(\mathbb{E}, \mathbb{E}^*)$ la convergence de X_{n_i} vers Z. (*ii*) $\Rightarrow$ (*iii*). Soit $Y \in \mathbb{E}^*_+$. L'hypothèse (*ii*) entraîne que $(|X_n|Y)$ est équi-intégrable ; il en est de même pour $(X_n^- Y)$ et $(X_n^+ Y)$ et on continue comme précédemment. (*iii*) $\Rightarrow$ (*i*) se traite de manière analogue.

Rappelons qu'une suite de v. a. (X_n) de $\mathbb{E}$ est dite $\mathbb{E}$-*équi-intégrable* si $\operatorname*{Sup}_n \| X_n 1_{\{|X_n| > a\}} \|_{\mathbb{E}} \to 0$ quand a tend vers l'infini. Il résulte aisément de ce qui précède.

COROLLAIRE 5. — *Soit* $\mathbb{E}$ *un espace de Köthe faiblement séquentiellement complet. Toute suite bornée,* $\mathbb{E}$ *équi-intégrable, est faiblement relativement compacte dans* $\mathbb{E}$.

Remarque. — Soit $\mathbb{E}$ un espace de Köthe n. c. o. et soit (X_n) une suite de v. a. de $\mathbb{E}$. On a les équivalences suivantes :

i) (X_n) est $\mathbb{E}$-équi-intégrable.

ii) Pour toute suite (h_n) vérifiant $\| h_n \|_\infty \leq 1$ et $\| h_n \| \to 0$ alors $\| X_n h_n \|_{\mathbb{E}} \to 0$.

ii') Pour toute suite (h_p) vérifiant $\| h_p \|_\infty \leq 1$ et $\| h_p \|_1 \to 0$ alors $\operatorname{Sup}_n \| X_n h_p \|_{\mathbb{E}} \xrightarrow[p \to \infty]{\mathbb{E}} 0$.

iii) Pour toute suite A_n vérifiant $P(A_n) \to 0$ $\| X_n 1_{A_n} \|_{\mathbb{E}} \xrightarrow[n \to \infty]{} 0$.

iii') Pour toute suite A_n vérifiant $P(A_n) \to 0$ $\operatorname{Sup}_n \| X_n 1_{A_p} \|_{\mathbb{E}} \xrightarrow[p \to \infty]{} 0$.

En effet on a tout d'abord (*i*) $\Leftrightarrow$ (*iii'*) en utilisant l'argument classique de L^1 (voir [*23*] proposition II.5.2). Il est clair que (*iii*) et (*iii'*) sont équivalents ainsi que (*ii*) et (*ii'*) enfin (*ii*) $\Rightarrow$ (*iii*) est trivial. Montrons (*iii*) $\Rightarrow$ (*ii*). Nous pouvons supposer que $h_n \to 0$ p. s., c'est-à-dire, par le théorème d'Egoroff $\| h_n 1_A \|_\infty \to 0$ pour des A de probabilités arbitrairement grandes, on a alors $\| X_n h_n \|_{\mathbb{E}} \leq \| X_n 1_{A^c} \|_{\mathbb{E}} + \| X_n \|_{\mathbb{E}} \cdot \| h_n 1_A \|_\infty$, aussi petit qu'on le souhaite.

Dans L^1 la faible compacité équivaut à l'équi-intégrabilité. Nous allons voir que cette situation est exceptionnelle.

COROLLAIRE 6. — *Soit* $\mathbb{E}$ *un Köthe faiblement séquentiellement complet tel que* L^∞ *soit fortement dense dans* $\mathbb{E}^*$. *Alors* $\mathbb{E}$ *est isomorphe à un espace* L^1 *si et seulement si « toute suite faiblement convergente est équi-intégrable dans l'espace ».*

Démonstration. — Soit (g_n) une suite de $\mathbb{E}_+$ convergeant faiblement vers 0. Alors $g_n 1_{\{g_n \leqslant a\}}$ converge faiblement vers 0 donc fortement (cf. [*11*]). De plus l'équi-intégrabilité assure que $\| g_n 1_{\{g_n \geqslant a\}} \| \leqslant \varepsilon$ si $a \geqslant A(\varepsilon)$ donc $\| g_n \| \to 0$. Soit maintenant (f_n) une suite bornée en norme dans $\mathbb{E}$ convergeant vers 0 fortement dans L^1. D'après le corollaire 3, $f_n \to 0$ pour $\sigma(\mathbb{E}, \mathbb{E}^*)$ et la suite $(g_n = | f_n |)$ est faiblement relativement compacte dans $\mathbb{E}$ d'après le corollaire 4. Donc (g_n) converge faiblement vers 0 d'où $\| g_n \| \to 0$ et ainsi (f_n) converge vers 0 dans $\mathbb{E}$.

Remarque. — Si L^∞ n'est pas fortement dense dans $\mathbb{E}^*$, par exemple si $\mathbb{E} = L \log L$, le corollaire est faux.

2. Existence des meilleurs approximants.

THÉORÈME 7. — *Soit* $\mathbb{E}$ *un espace de Köthe de v. a. définies sur un espace de probabilité* $(\Omega, \mathfrak{A}, P)$, *vérifiant l'une des deux propriétés suivantes :*

1) $\mathbb{E}$ *est faiblement séquentiellement complet* (*c'est le cas de* L^1) *ou* 2) $\mathbb{E}$ *possède la propriété de Fatou et ses parties bornées sont équi-intégrables dans* L^1 (*c'est le cas de* L^∞).

Alors, si $\mathscr{B}$ *est une sous-tribu de* $\mathfrak{A}$, *pour tout* $X \in \mathbb{E}$ *il existe au moins une v. a.* $Z \in \mathbb{E}(\mathscr{B})$ *vérifiant* $\|X-Z\|_{\mathbb{E}} \leqslant \|X-Y\|_{\mathbb{E}}$ *pour toute v. a.* $Y \in \mathbb{E}(\mathscr{B})$.

Démonstration. — Soit (X_n) une suite de v. a. de $\mathbb{E}(\mathscr{B})$ vérifiant

$$\|X - X_n\|_{\mathbb{E}} \to \inf \{ \|X - Y\|_{\mathbb{E}}; Y \in \mathbb{E}(\mathscr{B}) \}.$$

Nous disons que (X_n) est une suite approximante de X sachant $\mathscr{B}$ et nous notons $a(X_n) = \underline{\lim}\, E(|X - X_n|)$. Soit $a = \inf \{ a(X_n) ;$ pour toute suite approximante (X_n) de X sachant $\mathscr{B} \}$. Par le procédé diagonal nous construisons une suite approximante (X_n) de X telle que $E(|X - X_n|) \to a$. Montrons que cette suite (X_n) est équi-intégrable dans L^1. Soit (B_n) une suite d'événements de $\mathscr{B}$ vérifiant $P(B_n) \to 0$ et posons $Y_n = X_n 1_{B_n^c}$. On a alors $|X - X_n| 1_{B_n} + |X - Y_n| = |X| 1_{B_n} + |X - X_n|$. Supposons que $\mathbb{E}$ soit faiblement séquentiellement complet ; il est en particulier n. c. o. et donc on a $\|X 1_{B_n}\|_{\mathbb{E}} \to 0$. Comme la suite (Y_n) est manifestement approximante pour X sachant $\mathscr{B}$, on en déduit que $E(|X - X_n| 1_{B_n}) \to 0$. Enfin l'inégalité : $|X_n| 1_{B_n} \leqslant |X - X_n| 1_{B_n} + |X| 1_{B_n}$ montre que la suite (X_n) est équi-intégrable dans $L^1(\mathscr{B})$ donc dans L^1.

Supposons maintenant que $\mathbb{E}$ vérifie la propriété 2). Si (X_n) est une suite approximante de X sachant $\mathscr{B}$, elle est bornée dans $\mathbb{E}$ donc équi-intégrable dans L^1. Par suite nous pouvons supposer que X_n converge faiblement dans L^1 vers une v. a. X_0, ceci dans les deux hypothèses. Comme les v. a. X_n sont $\mathscr{B}$-mesurables il en est de même pour X_0. Enfin $\mathbb{E}$ possédant, dans les deux cas envisagés, la propriété de Fatou, le lemme 1 assure que X_0 appartient à $\mathbb{E}(\mathscr{B})$ et, pour toute v. a. $Y \in \mathbb{E}(\mathscr{B})$

$$\|X - X_0\|_{\mathbb{E}} \leqslant \underline{\lim}\, \|X - X_n\|_{\mathbb{E}} \leqslant \|X - Y\|_{\mathbb{E}}.$$

3. Propriétés des meilleurs approximants.

Soit $\mathbb{E}$ un espace de Köthe possédant l'une des propriétés du théorème précédent, par exemple un espace de Orlicz muni d'une norme habituelle. Notons $\mathscr{A}_{\mathscr{B}}(X)$ l'ensemble des meilleurs approximants d'une v. a. $X \in \mathbb{E}$

sachant une sous-tribu $\mathscr{B}$ de $\mathfrak{A}$. Les propriétés suivantes se vérifient aisément

i) Si $X \in \mathbb{E}_+$, $\mathscr{A}_{\mathscr{B}}(X) \subset \mathbb{E}_+(\mathscr{B})$.

ii) Si $B \in \mathscr{B}$ et si $Y \in \mathscr{A}_{\mathscr{B}}(X1_B)$ alors $Y1_{B^c} = 0$ p. s.

iii) Si $Y \in \mathbb{E}(\mathscr{B})$ $Y + \mathscr{A}_{\mathscr{B}}(X) = \mathscr{A}_{\mathscr{B}}(X + Y)$.

iv) $\mathscr{A}_{\mathscr{B}}(X)$ est un sous-ensemble convexe fermé de $\mathbb{E}$.

v) Si $\mathbb{E}$ est strictement convexe alors $\mathscr{A}_{\mathscr{B}}(X)$ est réduit à un élément.

vi) Si $X_n \to X$ dans $\mathbb{E}$ et si $Y_n \in \mathscr{A}_{\mathscr{B}}(X_n)$, pour tout n, la suite (Y_n) est approximante pour X, sachant $\mathscr{B}$.

A titre d'exemple, montrons cette dernière propriété :

$$\text{pour tout} \quad Y \in \mathbb{E}(\mathscr{B}) \qquad \text{on a} \quad \| Y_n - X_n \|_{\mathbb{E}} \leq \| Y - X_n \|_{\mathbb{E}} ;$$

il en résulte :

$$\| Y_n - X \|_{\mathbb{E}} \leq \| Y_n - X_n \|_{\mathbb{E}} + \| X_n - X \|_{\mathbb{E}} \leq \| X - Y \|_{\mathbb{E}} + 2 \| X - X_n \|_{\mathbb{E}},$$

d'où le résultat.

Remarques. — Si $\mathbb{E} = L^p$, $1 < p < \infty$, $\mathscr{A}_{\mathscr{B}}(X)$ est réduit à un seul élément, la p-moyenne conditionnelle, [*1*] [*2*]; notons la $E_p^{\mathscr{B}}(X)$. Si $p = 2$, on retrouve l'espérance conditionnelle $E^{\mathscr{B}}(X)$. Cette p-moyenne conditionnelle jouit d'une propriété remarquable, à savoir : pour tout $B \in \mathscr{B}$: $E_p^{\mathscr{B}}(X1_B) = 1_B E_p^{\mathscr{B}}(X)$ (cf. [*1*] et § 4).

Herrndorf a montré dans [*12*] que cette propriété était mise en défaut dans tout espace d'Orlicz muni de la norme de Luxemburg, différent d'un espace L^p, $1 < p < \infty$. Il n'est donc pas question qu'elle soit vérifiée dans le présent contexte (voir § 4).

L'espérance conditionnelle $E^{\mathscr{B}}(X)$ possède de plus la propriété de transitivité : si $\mathscr{B} \subset \mathscr{C} \subset \mathfrak{A}$ on a $E^{\mathscr{B}}(X) = E^{\mathscr{B}}(E^{\mathscr{C}}(X))$. Cette propriété qui correspond au théorème des 3 perpendiculaires est fausse déjà dans les espaces L^p, $p \neq 2$. De la même façon les meilleurs approximants ne sont linéaires que dans l'espace L^2, (cf. [*16*]) et croissants que dans les espaces L^p; $1 < p < \infty$. Toutes ces propriétés sont donc irrécupérables dans un cadre aussi général que celui présenté ici. Nous les retrouverons dans la deuxième partie.

La dernière propriété, notée *vi*) ci-dessus, suggère que les suites approximantes jouissent de propriétés intéressantes comme dans le cas des espaces L^p, (cf. [*16*]). C'est ce que nous voyons maintenant. Au préalable nous rappelons la définition suivante, (cf. [*4*] p. 371-373).

Un espace de Banach réticulé est *uniformément monotone* si pour tout $\varepsilon > 0$ il existe $\delta > 0$ tels que :

$$f, g \geqslant 0, \qquad \| f \| = 1, \qquad \| f + g \| < 1 + \delta \Rightarrow \| g \| \leqslant \varepsilon .$$

Tout espace de Banach réticulé uniformément convexe est uniformément monotone. Un treillis de Banach est uniformément monotone si et seulement si, pour tout $\varepsilon > 0$ il existe $\delta > 0$ tels que :

$$0 \leqslant g \leqslant f, \qquad \| f \| = 1, \qquad \| f - g \| > \varepsilon \Rightarrow \| g \| \leqslant 1 - \delta .$$

De plus un tel espace est nécessairement faiblement séquentiellement complet. On vérifie aisément que tout espace d'Orlicz modéré L^Φ muni de la norme de Luxemburg : $\| X \|_\Phi = \inf\left\{ k : E\left(\Phi\left(\frac{|X|}{k}\right)\right) \leqslant 1 \right\}$ est uniformément monotone, alors qu'il n'est pas, en général, uniformément convexe. Observons enfin que $\mathbb{E}$ est uniformément monotone, si (x_n) et (u_n) sont deux suites de $\mathbb{E}_+$ vérifiant $\lim \| u_n + x_n \| = \lim \| u_n \|$ alors, on a $\lim \| x_n \| = 0$.

Les suites approximantes des espaces de Köthe uniformément monotones possèdent les mêmes propriétés que celles des espaces L^p.

Proposition 8. — *Soit $\mathbb{E}$ un espace de Köthe uniformément monotone, de v. a. définies sur un espace de probabilité* $(\Omega, \mathfrak{A}, P)$ *et soit $\mathscr{B}$ une sous-tribu de $\mathfrak{A}$ et* X *une v. a. de $\mathbb{E}$. Alors toute suite approximante de* X *sachant $\mathscr{B}$ est $\mathbb{E}$-équi-intégrable et ses points d'accumulations pour $\sigma(\mathbb{E}, \mathbb{E}^*)$ sont des meilleurs approximants de* X *sachant $\mathscr{B}$. L'ensemble $\mathscr{A}_{\mathscr{B}}(X)$ est une partie convexe faiblement compacte de $\mathbb{E}$.*

Démonstration. — Soit (X_n) une suite approximante de X sachant $\mathscr{B}$ et soit B_n une suite d'événements de $\mathscr{B}$ tels que $P(B_n) \to 0$. Nous notons, comme dans la démonstration du théorème 7, $Y_n = X_n 1_{B_n^c}$; l'égalité $|X - X_n| 1_{B_n} + |X - Y_n| = |X| 1_{B_n} + |X - X_n|$ montre que Y_n est une suite approximante de X sachant $\mathscr{B}$ et que $\lim \| (X - X_n) 1_{B_n} \|_{\mathbb{E}} = 0$, $\mathbb{E}$ étant uniformément monotone. Par conséquent X_n est $\mathbb{E}(\mathscr{B})$-équi-intégrable, et d'après la remarque 1-*iii*) du corollaire 5, la suite (X_n) est $\mathbb{E}$-équi-intégrable. (X_n) est donc faiblement relativement compacte. Soit X_0 un point d'accumulation faible de (X_n). Comme $|X - X_n|$ est également faiblement relativement compacte (corollaire 4), il existe une sous-suite (n_i) telle que $X_{n_i} \to X_0$ et $|X - X_{n_i}| \to Z$ dans $\sigma(\mathbb{E}, \mathbb{E}^*)$ avec $Z \geqslant |X - X_0|$ il en résulte : $\| X - X_0 \|_{\mathbb{E}} \leq \| Z \|_{\mathbb{E}} \leq \underline{\lim} \, \| X - X_{n_i} \|_{\mathbb{E}}$ (lemme 1), ce qui achève la démonstration.

Corollaire 9. — *Soit $\mathbb{E}$ un espace de Köthe uniformément monotone vérifiant la propriété de Kadec : « si $X_n \to X$ faiblement et si $\| X_n \| \to \| X \|$ alors $\| X_n - X \| \to 0$ ». Alors de toute suite approximante de* X *sachant $\mathscr{B}$*

on peut extraire une sous-suite convergeant fortement vers un meilleur approximant de X *et l'ensemble* $\mathscr{A}_{\mathscr{B}}(X)$ *est alors une partie convexe compacte de* $\mathbb{E}$.

Remarque. — Supposons que $\mathscr{A}_{\mathscr{B}}(X)$ soit réduit à un seul point $p_{\mathscr{B}}(X)$. C'est le cas, nous l'avons dit, lorsque $\mathbb{E}$ est strictement convexe. Il résulte des énoncés précédents que :

Si $\mathbb{E}$ vérifie les hypothèses de la proposition 8, l'application $X \to p_{\mathscr{B}}(X)$ est faiblement continue dans $\mathbb{E}$.

Si $\mathbb{E}$ vérifie les hypothèses du corollaire 9, l'application $X \to p_{\mathscr{B}}(X)$ est fortement continue dans $\mathbb{E}$.

Herrndorf a obtenu un résultat semblable pour les espaces d'Orlicz (cf. [*13*] corollaire 3.19).

A titre d'illustration, nous allons examiner certaines propriétés particulières aux meilleurs approximants dans les espaces L^p, $1 \leqslant p \leqslant \infty$.

4. Cas des espaces L^p $1 \leqslant p \leqslant +\infty$.

Un espace de Köthe est un L^p, $1 \leqslant p \leqslant \infty$ si, pour tout choix $f, g, h \in \mathbb{E}_+$ avec $(f+g) \wedge h = 0$ et $\|f\| \leqslant \|g\|$ on a $\|f+h\| \leqslant \|g+h\|$ ([*21*] I-*b*-7). Soit maintenant $\mathscr{B}$ une sous-tribu de $\mathfrak{A}$ alors pour tout $B \in \mathscr{B}$ et tout $X \in L^p$ on a :

$$\mathscr{A}_{\mathscr{B}}(X1_B) + \mathscr{A}_{\mathscr{B}}(X1_{B^c}) \subset \mathscr{A}_{\mathscr{B}}(X)$$

et

$$\mathscr{A}_{\mathscr{B}}(X)1_B + \mathscr{A}_{\mathscr{B}}(X1_{B^c}) \subset \mathscr{A}_{\mathscr{B}}(X).$$

Pour cela posons

$$f_1 = |X1_B - p(X1_B)|\,; \qquad f_2 = |X1_B - p(X)1_B|\,;$$
$$g_1 = |X1_{B^c} - p(X1_{B^c})| \qquad \text{et} \qquad g_2 = |X1_{B^c} - p(X)1_{B^c}|$$

où $p(X)$ désigne un élément de $\mathscr{A}_{\mathscr{B}}(X)$. En considérant les triplets (f_1, f_2, g_1) et (g_1, g_2, f_2), les inégalités suivantes

$$\|X - (p(X1_B) + p(X1_{B^c}))\| \leqslant \|X - (p(X)1_B + p(X1_{B^c}))\| \leqslant \|X - p(X)\|$$

montrent les deux relations cherchées. De plus si $p \neq \infty$ l'égalité $\|f+h\| = \|h+g\|$ implique $\|f\| = \|g\|$ donc, pour $p \neq \infty$ nous avons $\mathscr{A}_{\mathscr{B}}(X)1_B \subset \mathscr{A}_{\mathscr{B}}(X1_B)$, ce qui redémontre les résultats de [*2*] p. 35 et de [*1*]. Dans le cas de L^∞ nous n'obtenons que $\mathscr{A}_{\mathscr{B}}(X)1_B \subset \mathscr{A}_{\mathscr{B}}(X1_B)$ ou $\mathscr{A}_{\mathscr{B}}(X)1_{B^c} \subset \mathscr{A}_{\mathscr{B}}(X1_{B^c})$.

Nous allons montrer qu'il existe dans L^1 et dans L^∞ un bon choix de meilleurs approximants, limites d'approximants dans les L^p, $1 < p < +\infty$.

PROPOSITION 10. — *Soit $\mathscr{B}$ une sous-tribu de $\mathfrak{A}$ et soit $X \in L^\infty$, les points d'accumulation de la suite $(E_n^{\mathscr{B}}(X))$ pour $\sigma(L^1, L^\infty)$ sont des meilleurs approximants de X dans L^∞, sachant $\mathscr{B}$. On a*

$$\lim_n \| X - E_n^{\mathscr{B}}(X) \|_\infty = \inf \{ \| X - Y \|_\infty ; Y \in L^\infty(\mathscr{B}) \}.$$

Démonstration. — Pour tout n on a $\| E_n^{\mathscr{B}}(X) \|_\infty \leqslant \| X \|_\infty$; on peut donc supposer par sous-suite, que $E_n^{\mathscr{B}}(X) \to X_0$ pour $\sigma(L^1, L^\infty)$. Soit $Y \in L^\infty(\mathscr{B})$.

L'inégalité $\| X - E_n^{\mathscr{B}}(X) \|_n \leqslant \| X - Y \|_n \leqslant \| X - Y \|_\infty$ montre que $\lim \| X - E_n^{\mathscr{B}}(X) \|_n \leqslant \inf \{ \| X - Y \| ; Y \in L^\infty(\mathscr{B}) \}$, la suite $\| X - E_n^{\mathscr{B}}(X) \|_n$ étant décroissante. Soit maintenant $Z \in L^1$ et pour $p > 0$, notons $Z \times p = -p \vee Z \wedge p$; il vient alors, à ε fixé et pour tout p et n assez grands

$$\begin{aligned} E[Z.(X - X_0)] &\leqslant \varepsilon + E[Z \times p.(X - X_0)] \leqslant 2\varepsilon + E[Z \times p.(X - E_n^{\mathscr{B}}(X))] \\ &\leqslant 2\varepsilon + \| Z \times p \|_m . \| X - E_n^{\mathscr{B}}(X) \|_n \quad \text{avec} \quad \frac{1}{m} + \frac{1}{n} = 1 \\ &\leqslant 2\varepsilon + \| Z \times p \|_1 . \lim \| X - E_n^{\mathscr{B}}(X) \|_n . \end{aligned}$$

En faisant tendre p vers l'infini, on obtient finalement

$$\| X - E_n^{\mathscr{B}}(X) \|_n \leqslant \| X - X_0 \|_\infty \leqslant \lim \| X - E_n^{\mathscr{B}}(X) \|_n \leqslant \inf \{ \| X - Y \| ; Y \in L^\infty(\mathscr{B}) \},$$

ce qui montre que X_0 est un meilleur approximant de X dans L^∞ sachant $\mathscr{B}$, et que $\lim \| X - E_n^{\mathscr{B}}(X) \|_n = \| X - X_0 \|_\infty$.

Remarques. — 1) De la même façon nous pouvons établir un résultat sur l'existence d'une médiane naturelle (cf. [*17*]). En effet soit $X \in L^1$ et $Y \in L^1(\mathscr{B})$; comme $X \times p$ tend vers X on a, pour ε donné, n assez grand et $\alpha > 1$, voisin de 1,

$$\varepsilon + \| X - Y \|_1 \geqslant \| X \times n - Y \times n \|_1 \geqslant \| X \times n - Y \times n \|_\alpha - \varepsilon$$

et $\| X - E_\alpha^{\mathscr{B}}(X \times n) \|_1 \geqslant \varepsilon + \| X \times n - E_\alpha^{\mathscr{B}}(X \times n \|_\alpha$ c'est-à-dire

$$\| X - E_\alpha^{\mathscr{B}}(X \times n) \|_1 \leqslant 3\varepsilon + \| X - Y \|_1 .$$

Il existe une suite (α_n) décroissant strictement vers 1, telle que

$$E_{\alpha_n}^{\mathscr{B}}(X \times n) \to E_1^{\mathscr{B}}(X) \qquad \text{pour} \qquad \sigma(L^1, L^\infty).$$

2) Soit $p(X)$ un meilleur approximant de X sachant $\mathscr{B}$ dans L^1 ou dans L^∞ obtenu par limite faible dans L^1 comme décrit dans ce qui précède. Il est alors clair que $1_B . p_{\mathscr{B}}(X) \in \mathscr{A}_{\mathscr{B}}(X 1_B)$. De la même façon, si $\mathscr{B}_n$ est une suite monotone de sous-tribus de $\mathfrak{A}$, par le procédé diagonal il est possible de

construire une suite (X_n) avec $X_n \in \mathscr{A}_{\mathscr{B}_n}(X)$ de meilleurs approximants de X dans L^1 ou L^∞, telle que si T est l'ensemble des temps d'arrêts bornés liés à la filtration $(\mathscr{B}_n)$ et si $\mathscr{B}_\tau$ est la tribu des événements antérieurs à τ, pour $\tau \in T$, on ait :

$$\sum_n X_n 1_{\{\tau=n\}} = X_\tau \in \mathscr{A}_{\mathscr{B}_\tau}(X) \qquad \text{(voir § 6)}.$$

5. Problèmes d'approximation dans les treillis de Banach n. c. o.

Soit $\mathbb{E}$ un treillis de Banach n. c. o. et $\mathbb{F}$ un sous-treillis de Banach. Soit $e \in \mathbb{E}$ on dit que $f_0 \in \mathbb{F}$ est un meilleur approximant de e dans $\mathbb{F}$ si pour tout $f \in \mathbb{F}$ $\| e - f_0 \| \leqslant \| e - f \|$. De même on appelle suite approximante de e dans $\mathbb{F}$ une suite (f_n) d'éléments de $\mathbb{F}$ vérifiant

$$\lim_n \| e - f_n \| = \inf \{ \| e - f \| ; f \in \mathbb{F} \}.$$

Il est alors possible de réduire le problème de la recherche d'un meilleur approximant en considérant les treillis de Banach $\mathbb{F}_0$ engendrés par (f_n) et, $\mathbb{E}_0$, engendré par (f_n) et e. On sait alors que $\mathbb{E}_0$ est séparable. Ainsi $\mathbb{E}_0$ possède une unité faible et on se trouve ramené à un espace de Köthe par le théorème d'isomorphisme (cf. [*21*]-I-*b*-14).

Soit $\mathbb{E}$ un espace de Köthe n. c. o. et $\mathbb{F}$ un sous Banach réticulé. Si $\mathbb{F}$ contient les constantes, il est facile de montrer (cf. [*23*] p. 36) que $\mathbb{F}$ est de la forme $\mathbb{E}(\mathscr{B})$ où $\mathscr{B}$ est une sous-tribu de $\mathfrak{A}$; on se trouve exactement dans la situation décrite au paragraphe précédent. Si $\mathbb{F}$ ne contient pas les constantes, on peut après réduction, se borner au cas $\mathbb{F}$ séparable. Soit f_0 une unité faible de $\mathbb{F}$; on montre facilement, en adaptant la construction classique, qu'il existe une sous-tribu $\mathscr{B}$ telle que

$$\mathbb{F} = \{ X \in \mathbb{E} ; X = Y f_0, Y\ \mathscr{B}\text{-mesurable} \}.$$

Soit (X_n) une suite approximante de X dans $\mathbb{F}$, nous pouvons écrire $X_n = Y_n f_0$; la démonstration du théorème 7 montre que $E(|X_n| 1_{B_n}) \to 0$ lorsque $P(B_n) \to 0$, $B_n \in \mathscr{B}$. Posons $Q = f_0 . P$. On peut extraire de la suite (Y_n) une sous-suite convergeant vers une v. a. Y pour $\sigma(L^1(\mathscr{B}, Q), L^\infty(\mathscr{B}, Q))$. Notons encore (Y_n) cette sous-suite. Soit $\phi \in L^\infty_+$ et posons $\overline{Q} = \phi . Q$. L'application canonique de $L^1(\mathscr{B}, Q)$ dans $L^1(\mathscr{B}, \overline{Q})$ est continue, donc la suite (Y_n) converge faiblement vers Y dans $L^1(\mathscr{B}, \overline{Q})$. Ainsi (X_n) converge faiblement dans $L^1(\mathfrak{A}, P)$ vers une v. a. Y_0, de la forme $Y f_0$, qui appartient donc à $\mathbb{F}$ et se trouve être un meilleur approximant de X. On a obtenu une version abstraite du théorème 7 :

THÉORÈME 7 *bis*. — *Soit* $\mathbb{E}$ *un treillis de Banach faiblement séquentiellement complet et* $\mathbb{F}$ *un sous-Banach réticulé de* $\mathbb{E}$. *Pour tout* $X \in \mathbb{E}$, *il existe* $Y_0 \in \mathbb{F}$ *tel que* $\| X - Y_0 \| \leqslant \| X - Y \|$ *pour tout* $Y \in \mathbb{F}$.

Considérons maintenant une filtration $(\mathscr{B}_n)$ croissante ou décroissante de sous-tribu de $\mathfrak{A}$, dans l'espace de probabilité $(\Omega, \mathfrak{A}, P)$. Amemiya et Ando ont montré dans [*1*] que pour $1 < p < +\infty$ et $X \in L^p$ la suite $E_p^{\mathscr{B}}(X)$ convergeait p. s. et dans L^p. Ce résultat a été généralisé par Ando et Shintani dans [*2*], au cas L^1 ; puis pour les espaces d'Orlicz muni de la norme de Luxemburg par Landers et Rogge, [*15*], voir aussi [*13*]. Nous examinons rapidement pour conclure cette première partie, ces questions de convergence.

6. Convergence des meilleurs approximants.

Soit $\mathbb{E}$ un espace de Köthe de v. a. r. définies sur un espace de probabilité $(\Omega, \mathfrak{A}, P)$ et satisfaisant l'une ou l'autre propriété du théorème 7. Soit $(\mathscr{B}_n)$ une filtration croissante vérifiant $\vee \mathscr{B}_n = \mathfrak{A}$, le cas décroissant se traiterait de façon analogue. Pour $X \in \mathbb{E}$, choisissons $X_n \in \mathscr{A}_{\mathscr{B}_n}(X)$, pour tout n. Supposons que $\bigcup_n \mathbb{E}(\mathscr{B}_n)$ soit dense dans $\mathbb{E}$ (c'est le cas si $\mathbb{E}$ est n. c. o.) il existe alors une suite $Y_n \in \mathbb{E}(\mathscr{B}_n)$ telle que $\| Y_n - X \| \to 0$ et on a : $\| X - X_n \| \leqslant \| X - Y_n \| \to 0$. Si donc $\mathbb{E}$ est n. c. o., les suites de meilleurs approximants suivant une filtration, convergent toujours pour la norme de $\mathbb{E}$ et donc en particulier, dans L^1 et en probabilité. Par contre, la convergence p. s. n'est en général pas assurée. Soit T l'ensemble des temps d'arrêt bornés liés à la filtration $(\mathscr{B}_n)$; pour $\tau \in T$, notons $\mathscr{B}_\tau$ la tribu des événements antérieurs à τ et $X_\tau = \sum_n X_n 1_{\{\tau = n\}}$. Si on a $X_\tau \in \mathscr{A}_{\mathscr{B}\tau}(X)$ pour tout $\tau \in T$, on a : $\| X - X_\tau \| \leqslant \| X - Y_n \|$ dès que $\tau \geqslant n$, p. s. et par conséquent $X_\tau \to X$ dans $\mathbb{E}$ donc en probabilité d'où $X_n \to X$ p. s. (cf. [*6*]). Cet argument ne s'applique seulement qu'aux cas des L^p pour $1 < p < +\infty$, ou des espaces d'Orlicz, L^U, muni de leur pseudo-norme $E(U(|X|))$; en effet la condition $X_\tau \in \mathscr{A}_{\mathscr{B}_\tau}(X)$ est alors satisfaite (cf. [*1*] [*10*]). Il ne semble pas possible de l'adapter à des espaces plus généraux et Herrndorf a montré que c'était impossible pour la norme de Luxemburg des espaces d'Orlicz différents des espaces L^p. Par contre, il est facile d'adapter ce qui précède au cas des fonctionnelles de types convexe, définies sur des espaces de Köthe. C'est ce que nous voyons maintenant.

Soit $\mathbb{E}$ un espace de Köthe de v. a. r. faiblement séquentiellement complet et soit U une fonction de $\mathbb{E}_+$ dans $\mathbb{R}_+$ vérifiant les propriétés suivantes :

i) $U(0) = 0$

ii) U est convexe sur $\mathbb{E}_+$ et $U(X) + U(Y) \leqslant U(X + Y)$ pour $X, Y \in \mathbb{E}_+$

iii) $U(|X_n|) \to 0$ si et seulement si $\|X_n\| \to 0$.

On vérifie alors que U est continue sur $\mathbb{E}_+$ et que la convergence faible, dans $\mathbb{E}_+$, de X_n vers X implique $U(X) \leqslant \underline{\lim}\ U(X_n)$.

Nous dirons que Y_0 est un *meilleur* U-*approximant* de X sachant une sous-tribu $\mathscr{B}$, si $Y_0 \in \mathbb{E}(\mathscr{B})$ et si $U(|X - Y_0|) \leqslant U(|X - Y|)$, tout $Y \in \mathbb{E}(\mathscr{B})$. Nous dirons que Y_n est une suite U-approximante de X sachant $\mathscr{B}$, si on a : $Y_n \in \mathbb{E}(\mathscr{B})$, pour tout n, et $\lim U(|X - Y_n|) = \inf U(|X - Y|) ; Y \in \mathbb{E}(\mathscr{B})$. On a en notant comme précédemment $\mathscr{A}_{\mathscr{B}}(X)$ l'ensemble des meilleurs U-approximants de X sachant $\mathscr{B}$.

THÉORÈME 11. — *Soit* $\mathbb{E}$ *un espace de Köthe faiblement séquentiellement complet et* U *une fonctionnelle vérifiant les propriétés* (*i*) (*ii*) *et* (*iii*) *ci-dessus,* $\mathscr{B}$ *une sous-tribu de* $\mathfrak{A}$. *Alors toute suite* U-*approximante de* X *sachant* $\mathscr{B}$ *est* $\mathbb{E}$-*équi-intégrable et ses points d'accumulation pour* $\sigma(\mathbb{E}, \mathbb{E}^*)$ *sont des* U-*meilleurs approximants de* X *sachant* $\mathscr{B}$. *Si en outre* U *vérifie* (*iv*) « *Si* $\inf(X, Y) = 0$ *pour* $X, Y \in \mathbb{E}_+$ *alors* $U(X + Y) = U(X) + U(Y)$ ». *Alors on a* $\mathscr{A}_{\mathscr{B}}(X) . 1_B \subset \mathscr{A}_{\mathscr{B}}(X1_B)$. *Enfin si* $(\mathscr{B}_n)$ *est une filtration croissante engendrant* $\mathfrak{A}$, *toute suite* $(X_n) \in \mathscr{A}_{\mathscr{B}_n}(X)$, *pour tout n, converge fortement et p. s. vers* X *dans* $\mathbb{E}$.

Démonstration. — Soit $B_n \in \mathscr{B}$ tels que $P(B_n) \to 0$ et soit (X_n) une suite U-approximante de X sachant $\mathscr{B}$; posons $Y_n = X_n = X_n 1_{B_n^c}$. La relation $|X - X_n 1_{B_n}| + |X - Y_n| = |X| 1_{B_n} + |X - X_n|$ et les propriétés (*i*), (*ii*), (*iii*) montrent que (Y_n) est une suite U-approximante et que $U(|X - X_n| 1_{B_n}) \to 0$. Il en résulte que (X_n) est $\mathbb{E}(\mathscr{B})$-équi-intégrable et donc $\mathbb{E}$-équi-intégrable, par l'argument donné dans la démonstration de la proposition 8. Le corollaire 5 montre que (X_n) est faiblement relativement compacte et le corollaire 4, qu'il en est de même de $|X - X_n|$. Soit X_0 un point d'accumulation faible ; pour une sous-suite convenable on a $X_{n_i} \to X_0$ et $|X_{n_i} - X| \to Z$ dans $\sigma(\mathbb{E}, \mathbb{E}^*)$ avec $Z \geqslant |X - X_0|$. Il en résulte : $U(|X - X_0|) \leqslant U(Z) \leqslant \underline{\lim}\ U(|X - X_n|)$ c'est-à-dire $X_0 \in \mathscr{A}_{\mathscr{B}}(X)$. Si U vérifie (*iv*) on montre comme au paragraphe 4, la relation demandée. Si $(\mathscr{B}_n)$ est une filtration et $X_n \in \mathscr{A}_{\mathscr{B}_n}(X)$ on a alors : $X_\tau \in \mathscr{A}_{\mathscr{B}_\tau}(X)$ pour $\tau \in T$.

Comme $\mathbb{E}$ est n. c. o., $\bigcup_n \mathbb{E}(\mathscr{B}_n)$ est dense dans $\mathbb{E}$. Il existe donc une suite $Z_n \in \mathbb{E}(\mathscr{B}_n)$ pour chaque n, convergeant vers X dans $\mathbb{E}$ comme $U(|X - X_\tau|) \leqslant U(|X - Z_n|)$ dès que $\tau \geqslant n$ p. s. il en résulte que $X_\tau \to X$ dans $\mathbb{E}$ et donc $X_n \to X$ p. s.

Remarques. — 1°) Le théorème 11 est valide sous des hypothèses plus faibles : (*i*) $U(0) = 0$. (*ii*) U est sur-additive et continue. (*iii*) $U(x_n) \to 0$ implique $x_n \to 0$ et $x_n \xrightarrow{\sigma} x$ implique $U(x) \leqslant \underline{\lim}\, U(x_n)$.

2°) Soit U une fonction d'Orlicz modérée, il est clair que la fonctionnelle $\tilde{U}(X) = E(U(|X|))$ définie sur l'espace L^U, vérifie les hypothèses du théorème 11. On retrouve ainsi certains résultats classiques.

3°) Supposons maintenant que $U : \mathbb{E} \to \mathbb{R}$ soit différentiable au sens de Gâteaux [*21*], c'est-à-dire que pour tout $X \in \mathbb{E}$ il existe $D_X \in \mathbb{E}^+$, telle que, $U(X + hY) = U(X) + hE(D_X . Y) + o(h)$ pour $h \in \mathbb{R}$. Une condition nécessaire pour qu'on ait $Y \in \mathscr{A}_{\mathscr{B}}(X)$, pour la fonctionnelle U, est que pour tout $Z \in \mathbb{E}(\mathscr{B})$, $D_{X-Y}(Z) = E(D_{X-Y} . Z) = 0$ c'est-à-dire $E^{\mathscr{B}}(D_{X-Y}) = 0$. Les meilleurs approximants s'expriment implicitement par l'intermédiaire de la loi conditionnelle.

Nous allons étudier plus en détails de tels approximants.

II. ψ-MÉDIANES CONDITIONNELLES

1. Introduction.

Soit $(\Omega, \mathfrak{A}, P)$ un espace probabilisé et soit ψ une fonction de $\mathbb{R}$ dans $\mathbb{R}$ croissante au sens large et nulle en 0. On note

$$M^\psi = \{ X \in L^0(\Omega, \mathfrak{A}, P) ; \text{il existe} \quad \alpha \in \mathbb{R} \quad \text{tel que} \quad \psi(X - \alpha) \in L^1 \}.$$

Le résultat suivant est aisé ([*14*] p. 64-65).

LEMME 1. — *Soit* $X \in M^\psi$ *et* $\lambda(\xi) = E[\psi(X - \xi)]$,

i) *$\lambda(\xi)$ existe pour tout ξ réel, en prenant éventuellement les valeurs $+ \infty$ ou $- \infty$*

ii) *λ est monotone décroissante, elle est strictement positive pour les grandes valeurs négatives de la variable et strictement négative pour ses grandes valeurs positives.*

Pour $X \in M^\psi$, nous noterons

$$\alpha(X) = \text{Sup} \{ \alpha ; E[\psi(X - \alpha)] > 0 \} \qquad \alpha'(X) = \text{Inf} \{ \alpha ; E[\psi(X - \alpha] < 0 \}.$$

Soit $\phi(x) = \int_0^x \psi(t)dt$. ϕ est convexe, positive, nulle en 0 et pour $X \in L^0$ nous posons $I = \{ \xi ; \psi(X - \xi) \in L^1 \}$ et $J = \{ \xi ; \Phi(X - \xi) \in L^1 \}$. I et J sont des intervalles ; en désignant par $\overset{\circ}{J}$ l'intérieur de J, on a :

LEMME 2. — *On a toujours* $\overset{\circ}{J} \subset I \subset J$.

Démonstration. — Soit $a \in \overset{\circ}{J}$; il existe $\varepsilon > 0$ tel que $[a - \varepsilon, a + \varepsilon] \subset J$. L'inégalité

$$(\varepsilon)^{-1}[\phi(X-a)-\phi(X-a-\varepsilon)] \leqslant \psi(X-a) \leqslant (\varepsilon)^{-1}[\phi(X-a+\varepsilon)-\phi(X-a)]$$

montre que $a \in I$. Montrons $I \subset J$. Il suffit de supposer que I n'est pas vide. Soit alors $a \in I$ et $c \in J$ et supposons par exemple $c \leqslant a$. Pour tout $b \in]c, a]$ on a $\psi(X - a) \leqslant \dfrac{\phi(X - c) - \phi(X - b)}{b - c}$ et comme ϕ est positive, on en déduit que $[a, c] \subset J$. D'où le résultat.

LEMME 3. — *Soit* $X \in M^{\psi}$, *les propriétés suivantes sont équivalentes :*

i) $\alpha(X) \leqslant \alpha \leqslant \alpha'(X)$

ii) $E[\phi(X - \alpha)] \leqslant E[\phi(X - a)]$ *pour tout a réel.*

Démonstration. — On a $\overset{\circ}{J} \subset I \subset J$ (lemme 2) soit $a \in J$ et vérifiant (*i*). L'intervalle $[\alpha, a]$ étant inclu dans J, on peut écrire

$$0 \leq \int_{\alpha}^{a} E[\psi(X - t)]dt = E(\Phi(X - a)) - E(\Phi(X - \alpha) \quad \text{et réciproquement.}$$

Soit maintenant $\mathscr{B}$ une sous-tribu de $\mathfrak{A}$. Pour $X \in L^0$ notons $F_X^{\mathscr{B}}(x) = E[1_{\{X \leqslant x\}}/\mathscr{B}]$ une version continue à droite de la fonction de répartition de X sachant $\mathscr{B}$. $F_X^{\mathscr{B}}(x)$ est p. s. une fonction de répartition et, pour tout x, $F_X(x)$ est $\mathscr{B}$-mesurable. Pour $X \in M^{\psi}$, la fonction

$$\alpha \to \int_{-\infty}^{+\infty} \psi(x - \alpha)dF_X^{\mathscr{B}}(x)$$

vérifie clairement les propriétés du lemme 1, p. s. On peut ainsi définir p. s. sans ambiguïté deux v. a. $\mathscr{B}$-mesurables par les formules :

$$\alpha_{\mathscr{B}}(X) = \operatorname{Sup}\left\{ \alpha ; \int_{-\infty}^{+\infty} \psi(x - \alpha)dF_X(x) > 0 \right\}$$

$$\alpha'_{\mathscr{B}}(X) = \operatorname{Inf}\left\{ \alpha ; \int_{-\infty}^{+\infty} \psi(x - \alpha)dF_X(x) < 0 \right\}$$

qui vérifient les propriétés : si Y est $\mathscr{B}$-mesurable, on a

$$\begin{aligned} Y \leq \alpha_{\mathscr{B}}(X) &\Leftrightarrow E(\psi(X - Y)/\mathscr{B}) \geq 0\,, \\ \alpha_{\mathscr{B}}(X) < Y < \alpha'_{\mathscr{B}}(X) &\Rightarrow E(\psi(X - Y)/\mathscr{B}) = 0\,, \\ Y \geq \alpha'_{\mathscr{B}}(X) &\Leftrightarrow E(\psi(X - Y)/\mathscr{B}) \leq 0\,. \end{aligned}$$

DÉFINITION. — Nous appellerons *ψ-médiane conditionnelle* de X sachant $\mathscr{B}$, toute v. a. Y $\mathscr{B}$-mesurable vérifiant : $\alpha_{\mathscr{B}}(X) \leqslant Y \leqslant \alpha'_{\mathscr{B}}(X)$. La proposition suivante est visiblement une conséquence du lemme 3 :

PROPOSITION 4. — *Si $X \in M^{\psi}$ et si Y est une v. a. $\mathscr{B}$-mesurable telle que $\alpha_{\mathscr{B}}(X) \leqslant Y \leqslant \alpha'_{\mathscr{B}}(X)$, alors $E[\phi(X - Y)] \leqslant E[\phi(X - Z)]$ pour toute v. a. Z $\mathscr{B}$-mesurable et réciproquement.*

Exemples. — 1°) Si $\psi(x) = x$, toute ψ-médiane conditionnelle Y de X sachant $\mathscr{B}$ vérifie : Y est $\mathscr{B}$-mesurable et $E(X - Y/\mathscr{B}) = 0$, c'est-à-dire $Y = E(X/\mathscr{B})$.

2°) Si $\psi(x) = p\,|\,x\,|^{p-1}\,\mathrm{sg}\,(x)$ avec $1 < p$, $\phi(x) = |\,x\,|^{p}$ on retrouve les p-moyennes conditionnelles d'Amemiya et Ando [*1*].

3°) $\psi(x) = \mathrm{sg}\,(x)$, $\phi(x) = |\,x\,|$, on retrouve les médianes conditionnelles d'Ando et Shintani [*2*], et de Laplace [*18*].

4°) Si $\psi(x) = -31_{\{x<0\}} + 1_{\{x>0\}}$, on obtient les premiers quartiles conditionnels, qui se trouvent ainsi résoudre un problème de meilleure approximation, à savoir celui relatif à la fonction $\phi(x) = x1_{\{x>0\}} - 3\,.\,x\,.\,1_{\{x<0\}}$. De la même façon tous les quantiles conditionnels résolvent un problème de meilleure approximation.

5°) Soit $X_1, \ldots, X_n, \ldots$ une suite de variables aléatoires indépendantes de même loi définie sur un espace de probabilité $(\Omega, \mathfrak{A}, P)$. Soit $\mathscr{S}_n$ la tribu des événements dépendants symétriquement de $X_1, \ldots, X_n$. Soit, enfin, ψ une fonction vérifiant les propriétés du début du paragraphe et ϕ sa primitive nulle en 0. Pour $X_1 \in M^{\psi}$ on a $E[\psi(X_1 - \alpha)/\mathscr{S}_n] = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} \psi(X_i - \alpha)$.

Si une v. a. $\mathscr{S}_n$-mesurable, Y, vérifie $\alpha_{\mathscr{S}_n}(X_1) \leqslant Y \leq \alpha'_{\mathscr{S}_n}(X_1)$ alors Y est un M-estimateur du paramètre de translation de la loi de X_1 (cf. [*14*]), il minimise $E[\phi(X_1 - Y)]$ parmi toutes les v. a. $\mathscr{S}_n$-mesurables Y. D'autre part si la loi commune des X_i a pour densité, à une constante près, $e^{-\phi(x-\theta)}$ sur $\mathbb{R}$, θ étant un paramètre réel, Y est également un estimateur du maximum de vraisemblance de θ après avoir observé $(X_1, \ldots, X_n)$. La loi en $e^{-\phi(x)}$ joue ici le rôle dévolu à la loi de Gauss en e^{-x^2}, dans le cas $\psi(x) = x$. Cette remarque réconcilie le critère local du maximum de vraisemblance que Daniel Bernoulli adopta en 1777, [*3*], et le critère global de Laplace que ce dernier soutint la même année devant l'Académie des Sciences et que nous avons rappelé plus haut, [*19*].

6°) Si $\psi(x)$ est impaire, on retrouve, dans le cas particulier des sous-tribus et des fonctions Φ convexes, les Φ-approximants de Landers et Rogge [*15*]. Notons que, dans ce cas, l'ensemble des Φ-approximants sachant $\mathscr{B}$ est toujours un intervalle d'ordre, ce qui permet de préciser un peu le résultat très général de [*15*] théorème 14.

2. Propriétés des ψ-médianes conditionnelles.

Dans tout ce paragraphe on se donne comme précédemment un espace de probabilité $(\Omega, \mathfrak{A}, P)$, une sous-tribu $\mathscr{B}$ de $\mathfrak{A}$ et une fonction croissante ψ nulle en 0. Pour $X \in M^\psi$, on a :

a) Si ψ est impaire, $\alpha_{\mathscr{B}}(-X) = -\alpha'_{\mathscr{B}}(X)$.

b) Si $X_1 \leq X_2$ sont deux v. a. de M^ψ on a :

$$\alpha_{\mathscr{B}}(X_1) \leq \alpha_{\mathscr{B}}(X_2) \qquad \text{et} \qquad \alpha'_{\mathscr{B}}(X_1) \leq \alpha'_{\mathscr{B}}(X_2)\,.$$

De *a*) et *b*) qui sont triviales, il résulte :

c) Si ψ est impaire et si Y est une ψ-médiane conditionnelle de X sachant $\mathscr{B}$ on a toujours $|Y| \leq \alpha'_{\mathscr{B}}(|X|)$.

d) La définition même montre que pour $B \in \mathscr{B}$, $\alpha_{\mathscr{B}}(X)1_B = \alpha_{\mathscr{B}}(X1_B)$ et $\alpha'_{\mathscr{B}}(X)1_B = \alpha'_{\mathscr{B}}(X1_B)$. En particulier, si Y est une ψ-médiane conditionnelle de X sachant $\mathscr{B}$, $Y1_B$ est une ψ-médiane conditionnelle de $X1_B$ sachant $\mathscr{B}$.

e) Si ψ est strictement croissante, l'application $\zeta \to E[\psi(X-\zeta)]$ est strictement décroissante lorsqu'elle est finie. On a donc $\alpha_{\mathscr{B}}(X) = \alpha'_{\mathscr{B}}(X)$.

Il n'existe alors qu'une seule ψ-médiane conditionnelle de X sachant $\mathscr{B}$. Nous la noterons toujours, dans ce cas, $m^\psi_{\mathscr{B}}(X)$ ou plus simplement $m_{\mathscr{B}}(X)$, s'il n'y a pas de confusion possible : $m_{\mathscr{B}}(X)$ est un opérateur positif, non linéaire en général, de M^ψ dans $M^\psi(\mathscr{B})$.

Comme les espérances conditionnelles les ψ-médianes conditionnelles possèdent des propriétés de continuité monotone :

f) Si ψ est continue et strictement croissante, *$m_{\mathscr{B}}$ est continue pour l'ordre de* M^ψ, c'est-à-dire : si X_n converge p. s. vers X et si $\sup_n X_n$ et $\inf_n X_n$ appartiennent à M^ψ alors $m_{\mathscr{B}}(X_n)$ converge vers $m_{\mathscr{B}}(X)$ p. s. et, de plus, $\sup_n m_{\mathscr{B}}(X_n)$ et $\inf_n m_{\mathscr{B}}(X_n)$ appartiennent à M^ψ.

Démonstration. — Il suffit de montrer que, par exemple, si X_n décroît vers X alors il en est de même pour $m_{\mathscr{B}}(X_n)$ vers $m_{\mathscr{B}}(X)$. Notons Y la limite de la suite décroissante $m_{\mathscr{B}}(X_n)$; Y est $\mathscr{B}$-mesurable et vérifie $Y \geqslant m_{\mathscr{B}}(X)$ donc $E[\psi(X_n - Y)/\mathscr{B}] \geqslant 0$ p. s. et, comme $\psi(X_n - Y)$ décroît vers $\psi(X - Y)$ car ψ est continue, il en résulte $E[\psi(X - Y)/\mathscr{B}] = 0$ donc $Y = m_{\mathscr{B}}(X)$.

Remarque. — Cette propriété qui est analogue au lemme 7-3 (*i*) de [*16*] ne peut être améliorée. En effet, dans le cas classique où $\psi(x) = x$, on sait que les conditions $\sup_n X_n$ et $\inf_n X_n \in L^1$ ne peuvent être affaiblies, voir [*5*].

Nous allons montrer maintenant, en toute généralité, le théorème de convergence p. s. des ψ-médianes conditionnelles suivant une filtration convergente. Ce résultat rassemble le théorème des martingales équi-intégrables, le théorème de consistance des M-estimateurs et le théorème de convergence p. s. des meilleurs approximants de [*1*] et [*2*].

THÉORÈME 5. — *Soit $(\mathscr{F}_n)$ une filtration croissante ou décroissante vers une sous-tribu $\mathscr{F}_\infty$ de $\mathfrak{A}$ et soit ψ une fonction croissante nulle en* 0. *Pour* $X \in M^\psi$ *on a toujours*

$$\alpha_{\mathscr{F}_\infty}(X) \leqslant \liminf \alpha_{\mathscr{F}_n}(X) \leqslant \limsup \alpha'_{\mathscr{F}_n}(X) \leqslant \alpha'_{\mathscr{F}_\infty}(X)\,.$$

Démonstration. — Pour chaque $q \in Q$, choisissons des versions X_q de $E[\psi(X - q)/\mathscr{F}_\infty]$ et $X_{n,q}$ de $E[\psi(X - q)/\mathscr{F}_n]$, $n \in \mathbb{N}$. Il existe un ensemble négligeable N tel que pour $\omega \notin N$, $X_{n,q}(\omega) \to X_q(\omega)$. On a clairement $\alpha'_n(\omega) = \operatorname{Inf}\{ q \,;\, X_{n,q}(\omega) < 0 \}$ et $\alpha'(\omega) = \operatorname{Inf}\{ q \,;\, X_q(\omega) < 0 \}$. Nous allons montrer que $\overline{\lim}\, \alpha'_n(\omega) \leqq \alpha'(\omega)$ p. s. Fixons un $\omega \notin N$ et soit $\beta < \overline{\lim}\, \alpha'_n(\omega)$. Pour ε assez petit, soit (n_i) une sous-suite telle que $\beta + \varepsilon < \alpha'_{n_i}(\omega)$ il existe $q_{n_i} \in Q$ tel que $\alpha'_{n_i}(\omega) - \dfrac{\varepsilon}{2} < q_{n_i} < \alpha'_{n_i}(\omega)$ tout i, on a alors $\beta < q_{n_i}$ et $X_{n_i, q_{n_i}} \geqq 0$ d'où il résulte $X_{n_i,\beta} \geqq 0$ et $X_\beta \geqq 0$ c'est-à-dire $\beta < \alpha'(\omega)$. On montre la seconde inégalité de la même façon.

Le théorème 5 montre seulement des convergences p. s. Pour la convergence en moyenne il faut, comme on le sait, vérifier une condition d'équi-intégrabilité ; c'est ce que nous ferons au paragraphe suivant. Au préalable montrons un résultat plus faible.

PROPOSITION 6. — *Soit $(\mathscr{F}_n)$ une filtration croissante ou décroissante vers une sous-tribu $\mathscr{F}_\infty$ de $\mathfrak{A}$ et soit ψ une fonction croissante nulle en* 0 *de primitive ϕ nulle en* 0 *et modérée. Si (X_n) est une suite de meilleurs approximants de* X *sachant $\mathscr{F}_n$, dans L^ϕ, on a $X_n \to X$ p. s. et dans L^ϕ.*

Démonstration. — On note T l'ensemble des temps d'arrêt bornés liés à $(\mathscr{F}_n)$ et, comme en I 6, on montre que $E\Phi(X_\tau - X) \to 0$ suivant T d'où il résulte $X_\tau \to X$ p. s. et dans L^Φ.

Remarques. — 1) Si $(\mathscr{F}_n)$ est une filtration décroissante, il est clair que la condition de modération est inutile.

2) Comme ϕ est convexe, $X_\tau \to X$ dans L^ϕ donc dans L^1 ; la suite est donc un amart équi-intégrable. Nous verrons qu'en outre (X_n) est contrôlée sur toute sa trajectoire par une martingale.

3. Contrôle des ψ-médianes conditionnelles.

Pour ne pas alourdir les énoncés, nous supposerons que ψ est impaire de sorte que ϕ est paire. Nous notons L^ϕ l'espace d'Orlicz usuel. Tout d'abord :

PROPOSITION 7. — *Soit $\mathscr{B}$ une sous-tribu de $\mathfrak{A}$ et soit $X \in M^\psi$ tel que $\phi(X) \in L^1$. Si Y est une ψ-médiane conditionnelle de X sachant $\mathscr{B}$, on a* $\phi\left(\frac{Y}{2}\right) \leqq E(\phi(X)/\mathscr{B})$.

Démonstration. — Pour $a > 0$, $Y . 1_{\{Y > a\}}$ est une ψ-médiane conditionnelle de $X1_{\{Y > a\}}$ (propriété *d*) il vient $E[\psi(X - a)1_{\{Y > a\}}/\mathscr{B}] \geq 0$. Comme $\phi(X) \in L^1$ on peut intégrer cette inégalité sur l'intervalle $[0, Y]$ par rapport à la mesure de Lebesgue (lemme 2) ; il vient alors $E(\phi(X - Y)/\mathscr{B}) \leqq E(\phi(X)/\mathscr{B})$. Comme ϕ est convexe et paire, on a en utilisant le lemme 8 ci-dessous

$$2E\left(\Phi\left(\frac{Y}{2}\right)/\mathscr{B}\right) - E(\Phi(X)/\mathscr{B}) \leq E(\Phi(Y - X/\mathscr{B} \leq E(\Phi(X)/\mathscr{B})$$

d'où le résultat.

LEMME 8. — *Si ϕ est convexe alors* $2\phi\left(\frac{a}{2}\right) - \phi(b) \leqq \phi(a-b) \leqq \phi(a) - 2\phi\left(\frac{b}{2}\right)$ *et si ϕ est concave :* $\phi(a) - 2\phi\left(\frac{b}{2}\right) \leqq \phi(a - b) \leqq 2\phi\left(\frac{a}{2}\right) - \phi(b)$.

Remarque. — Si $\phi(x) = x^p$ les inégalités du lemme 8 s'écrivent $\frac{1}{2^{p-1}} a^p - b^p \leqq |a - b|^p \operatorname{sg}(a - b) \leqq a^p - \frac{1}{2^{p-1}} . b^p$ lorsque $1 < p < \infty$. $a^p - 2^{1-p}b^p \leqq |a - b|^p \operatorname{sg}(a - b) \leqq 2^{1-p}a^p - b^p$ pour $0 < p < 1$. Ce sont les inégalités de [*1*] et de [*16*] lemme 7-2. Nous montrerons des inégalités plus générales au paragraphe suivant.

Donnons un exemple de la proposition 7 : soit $\psi(x) = \operatorname{sg}(x)$, $\phi(x) = x$, on obtient $|Y| \leqq 2E(|X|/\mathscr{B})$. On retrouve le lemme 1 de [2] : la médiane est inférieure à deux fois la moyenne pour une variable $\geqq 0$. Il résulte clairement de la proposition précédente que, pour $X \in L^{\Phi}$, la famille des ψ-médianes conditionnelles de X est L^{ϕ} équi-intégrable. Nous allons étendre ce résultat aux variables de M^{ψ}.

LEMME 9. — *Soit ϕ croissante, impaire de $\mathbb{R}$ dans $\mathbb{R}$, on a toujours :*

$$\phi\left(\frac{a}{2}\right) - 2\phi(b) \leq \phi(a-b) \leq 2\phi(a) - \phi\left(\frac{b}{2}\right) \qquad \textit{pour tout} \quad a, b \geq 0\,.$$

Démonstration. — Soit $b \leqq a$, on a $\phi(a-b) \leqq \phi(a) \leqq 2\phi(a) - \phi\left(\frac{b}{2}\right)$ de plus $\phi\left(\frac{a}{2}\right) \leqq \phi(b) \vee \phi(a-b) \leqq \phi(b) + \phi(a-b)$ donc

$$\phi\left(\frac{a}{2}\right) - 2\phi(b) \leqq \phi(a-b)\,.$$

Nous en déduisons l'analogue de la proposition 7 :

PROPOSITION 10. — *Soit $\mathscr{B}$ une sous-tribu de $\mathfrak{A}$, X une v. a. telle que $\psi(\|X\|) \in L^1$ et Y une ψ-médiane conditionnelle de X sachant $\mathscr{B}$. On a*
$\left|\psi\left(\frac{Y}{2}\right)\right| \leqq 2.E[|\psi(X)|/\mathscr{B}]$. □

Démonstration. — Supposons d'abord $X \geqq 0$. Soit $a > 0$; de l'inégalité $E[\psi(X-a)1_{\{Y>a\}}/\mathscr{B}] \geq 0$ et du lemme 9 il résulte

$$E[\psi(X)/\mathscr{B}]1_{\{Y>a\}} \geq \frac{1}{2}\psi\left(\frac{a}{2}\right)1_{\{Y>a\}}\,.$$

Par conséquent $\frac{1}{2}\psi\left(\frac{Y}{2}\right) > b$ implique $E[\psi(X)/\mathscr{B}] \geqq b$ pour tout b de la forme $b = \frac{1}{2}\psi\left(\frac{c}{2}\right)$ où $c > 0$. On en déduit $E[\psi(X)/\mathscr{B}] \geqq \frac{1}{2}\psi\left(\frac{Y}{2}\right)$. Si X est de signe quelconque, on utilise la propriété (*c*) de (2) pour conclure.

COROLLAIRE 11. — *Soit* X *une v. a. telle que* $\psi(2X) \in L^1$, *la famille* $\mathscr{M}^{\psi}(X)$ *des ψ-médianes conditionnelles de* X *pour toutes les sous-tribus de* $\mathfrak{A}$ *est « L^{ψ}-équi-intégrable » c'est-à-dire* $\{\psi(Y) ; Y \in \mathscr{M}^{\psi}(X)\}$ *est équi-intégrable.*

On peut, maintenant, énoncer le résultat de convergence en moyenne annoncé à la fin du paragraphe général.

THÉORÈME 12. — *Soit $(\mathscr{F}_n)$ une filtration croissante ou décroissante, convergeant vers une tribu $\mathscr{F}_\infty$ et soit* X *une v. a. telle que* $\alpha_{\mathscr{F}_\infty}(X) = \alpha'_{\mathscr{F}_\infty}(X)$.

Alors si ψ est continue à l'origine et si $\psi(2X)$ appartient à L^1, pour toute suite (X_n) de ψ-médianes conditionnelles de X le long de la filtration $(\mathscr{F}_n)$, on a $\psi(X_n - X) \to 0$ dans L^1.

Démonstration. — Le corollaire 11 et le lemme 9 montrent que la suite $\psi(X_n - X)$ est équi-intégrable et comme X_n converge vers X p. s. avec la proposition 6, le théorème en résulte.

Remarque. — La proposition 10 montre que les suites (X_n) de ψ-médianes de X le long d'une filtration $(\mathscr{F}_n)$ sont « ψ-contrôlées » par la martingale $E(|\psi(X)|/\mathscr{F}_n)$. Inversement, la martingale précédente est contrôlée par la suite (X_n) de la façon suivante :

PROPOSITION 13. — *Soit $\mathscr{B}$ une sous-tribu de $\mathfrak{A}$ et X une v. a. positive vérifiant $\psi(X) \in L^1$. Si Y est une ψ-médiane conditionnelle de X sachant $\mathscr{B}$; alors* $E\left(\psi\left(\frac{X}{2}\right)/\mathscr{B}\right) \leqq 2\psi(Y)$.

Démonstration. — Supposons $\alpha_{\mathscr{B}}(X) < Y < \alpha'_{\mathscr{B}}(X)$; alors on a $E(\psi(X - Y)/\mathscr{B}) = 0$ et du lemme 9 on déduit $E\left(\psi\left(\frac{X}{2}\right)/\mathscr{B}\right) \leqq 2\psi(Y)$. Le cas général s'en déduit aisément.

Les propositions 10 et 13 montrent que les suites de ψ-médianes d'une variable positive ont un comportement très voisin de celui des martingales équi-intégrables. Nous allons préciser cette analogie.

4. Inégalités maximales des suites de ψ-médianes conditionnelles.

Établissons auparavant quelques résultats techniques sur certaines fonctions. Soit ϕ une fonction croissante, nulle en 0, de $\mathbb{R}_+$ dans $\mathbb{R}_+$, nous disons que ϕ est modérée si $\sup_x \frac{\phi(2x)}{\phi(x)} < \infty$ $\left(\text{avec la convention } \frac{0}{0} = 0\right)$. Nous dirons que ϕ est comodérée s'il existe a et q, $q > 1$, tels que $\phi(ax) \geqq q\phi(x)$ tout $x \geqq 0$, cf. [7].

LEMME 14. — *i) ϕ est modérée si et seulement si, il existe une constante $c > 0$ telle que, pour tout $(x, y) \in \mathbb{R}_+^2$, on ait :*

$$\phi(x + y) \leqq c(\phi(x) + \phi(y)).$$

ii) ϕ est comodérée si et seulement si, il existe une constante $\alpha > 0$ telle que, pour tout $(x, y) \in \mathbb{R}_+^2$, on ait :

$$\phi(x) + \phi(y) \leqq \phi(\alpha(x + y)).$$

Démonstration. — Si ϕ est modérée on a, pour une constante c, $\phi(x+y) \leqq c\phi\left(\frac{x+y}{2}\right) \leqq c(\phi(x)+\phi(y))$; la réciproque est triviale, il suffit de faire $x = y$. Si ϕ est comodérée on a, pour une constante $\alpha > 1$: $\phi(x) + \phi(y) \leqq 2\phi(x + y) \leqq \phi(\alpha(x + y))$, réciproquement en faisant $x = y$ il vient : $2\phi(x) \leqq \phi(2\alpha x) \leqq \phi((2 \vee 2\alpha)x)$. Prolongeons ϕ en posant $\phi(x) = -\phi(-x)$ pour $x < 0$, nous avons :

LEMME 15. — *i) ϕ est modérée si et seulement si, il existe deux constantes α et $\beta > 0$ telles que, pour tout $(a, b) \in \mathbb{R}_+^2$ on ait :*

$$\beta\phi(a) - \alpha\phi(b) \leqq \phi(a - b) \leqq \alpha\phi(a) - \beta\phi(b).$$

ii) ϕ est comodérée si et seulement si, il existe $\beta > 1$ et $\alpha, \gamma > 0$ tels que, pour tout $(a, b) \in \mathbb{R}_+^2$ on ait :

$$\beta\phi\left(\frac{a}{\gamma}\right) - \phi(\alpha b) \leqq \phi(a - b) \leqq \phi(\alpha a) - \beta\phi\left(\frac{b}{\gamma}\right).$$

Démonstration. — Il suffit de supposer $a \geqq b$, les premières inégalités s'écrivent alors, en posant $a - b = h$

$$\beta\phi(b + h) - \alpha\phi(b) \leq \phi(h) \leq \alpha\phi(b + h) - \beta\phi(b)$$

l'inégalité de droite est valable dès que $\alpha - \beta \geqq 1$, l'inégalité de gauche est équivalente à la comodération de ϕ d'après le lemme 14. Les secondes inégalités s'écrivent

$$\beta\phi\left(\frac{b + h}{\gamma}\right) - \phi(\alpha b) \leq \phi(h) \leq \phi\left(\alpha(b + h)\right) - \beta\phi\left(\frac{b}{\gamma}\right)$$

et elles sont visiblement impliquées par la comodération. Réciproquement en posant $h = b$ on obtient :

$$\phi(b)+\beta\phi\left(\frac{b}{\gamma}\right) \leq \phi(2\alpha b), \quad \text{soit :} \quad \phi(b)(1+\beta) \leq \phi(2\alpha\gamma b) \leq \phi((2\alpha\gamma \vee 1)b).$$

Remarquons tout de suite que la proposition 13 admet comme version lorsque ψ est modérée, et pour toute ψ-médiane, Y, de X sachant $\mathscr{B}$

$$k_1 \mathrm{E}(\psi(\mathrm{X})/\mathscr{B}) \leqq \psi(\mathrm{Y}) \leqq k_2 \mathrm{E}(\psi(\mathrm{X})/\mathscr{B}).$$

Donnons nous, sur un espace de probabilité $(\Omega, \mathscr{F}, \mathrm{P})$, une filtration croissante $(\mathscr{F}_n)$ vérifiant $\mathscr{F}_0 = \{ \emptyset, \Omega)$ et $\bigvee_n \mathscr{F}_n = \mathscr{F}$. Comme dans l'introduction soit ψ une fonction de $\mathbb{R}$ dans $\mathbb{R}$, croissante et nulle en 0. Pour $X \in M^\psi$ soit (X_n) une suite de ψ-médianes de X le long de la filtration; on note $X^* = \sup_n |X_n|$. Nous nous proposons d'étudier l'intégrabilité comparée de X et de X*. Landers et Rogge ont montré dans [*16*] que l'inégalité des sous-martingales positives s'étendait au cas des suites de meilleurs approximants d'une v. a. de L^p_+, $1 < p < \infty$. Adaptons leur méthode :

LEMME 16. — *Soit* $X \in M^\psi_+$ *et* (X_n) *une suite de* ψ*-médianes de* X *le long de la filtration* $(\mathscr{F}_n)$. *On a* $\int_{\{X^* > a\}} \psi(X - a)dP \geq 0$ *pour tout* $a > 0$.

Démonstration. — Posons $\tau = \inf \{ n \,; X_n > a \}$ où $\inf \emptyset = 0$, on a alors : $a1_{\{\tau = k\}} < X_k 1_{\{\tau = k\}}$ et par conséquent

$$\int_{\{\tau = k\}} \psi(X - a)dP \geq 0 . \qquad \text{D'où le résultat.}$$

De ce lemme résulte aisément l'inégalité de Doob pour les suites de ψ-médianes d'une v. a. $X \in L^\phi$. Soit ψ une fonction continue, strictement croissante, impaire ; alors, en notant $m_{\mathscr{F}_n}(X)$ la ψ-médiane conditionnelle de X sachant $\mathscr{F}_n$, nous avons :

LEMME 17. — « *Inégalité de Doob* ». *Soit* $X \in L^\phi$ *et* $X^* = \sup_n |m_{\mathscr{F}_n}(X)|$. *On a* $X^* \in L^\phi$ *et* $E\left[\phi\left(\frac{X^*}{2}\right)\right] \leqq E(\phi(X))$.

Démonstration. — On suppose tout d'abord que X est positive et bornée ; en intégrant l'inégalité maximale du lemme 16, par rapport à la mesure de Lebesgue sur R_+, il vient $E\Phi(X^* - X) \leq E\Phi(X)$ (Φ est paire) et comme Φ est convexe : $\Phi(X^* - X) \geq 2\phi\left(\frac{X^*}{2}\right) - \Phi(X)$, l'inégalité du lemme en résulte. Si maintenant on a $X \in L^\phi_+$, comme $m_{\mathscr{F}_n}(X \wedge k)$ converge p. s., pour n fixé, vers $m_{\mathscr{F}_n}(X)$ quand k tend vers l'infini, alors $(X \wedge k)^*$ converge p. s. vers X*, et l'inégalité est encore valide. Enfin pour $X \in L^\phi$, il suffit d'observer qu'on a $|m_{\mathscr{F}_n}(X)| \leqq m_{\mathscr{F}_n}(|X|)$ (propriété (c) de (2) pour conclure. Donnons ci-dessous une forme un peu plus générale de l'inégalité maximale, qui se déduit aisément des lemmes 8, 9, 15 et 16.

THÉORÈME 18. — « *Inégalités maximales* ». *Soit* ψ *une fonction croissante* *impaire*, $X \in M_+^\psi$ *et soit* (X_n) *une suite de* ψ*-médianes conditionnelles de* X *suivant la filtration* $(\mathscr{F}_n)$. *On a, pour tout* $a > 0$:

. $\psi\left(\frac{a}{2}\right)P\{X^* > a\} \leq 2\int_{\{X^*>a\}} \psi(X)dP.$

. *Si* ψ *est modérée, avec les notations du lemme 9*

$$\psi(a)P\{X^* > a\} \leq \frac{\alpha}{\beta}\int_{\{X^*>a\}} \psi(X)dP.$$

. *Si* ψ *est comodérée, il existe deux constantes* α *et* $\beta > 1$ *telles que*

$$\psi(a)P\{X^* > a\} \leq \frac{1}{\beta}\int_{\{X^*>a\}} \psi(\alpha X)dP.$$

. *Si* ψ *est convexe* : $2\psi\left(\frac{a}{2}\right)P\{X^* > a\} \leq \int_{\{X^*>a\}} \psi(X)dP$

. *Si* ψ *est concave* : $\psi(a)P\{X^* > a\} \leq 2\int_{\{X^*>a\}} \psi\left(\frac{X}{2}\right)dP.$

Il est possible d'obtenir de façon analogue des inégalités maximales inverses, dans le cas des filtrations régulières. Pour cela nous utilisons la méthode développée par Jui Lin Long dans [*22*]. Rappelons qu'une filtration $(\mathscr{F}_n)$ est régulière s'il existe une constante d telle que, pour tout $B \in \mathscr{F}$: $E(1_B/\mathscr{F}_n) \leq d \,.\, E(1_B/\mathscr{F}_{n-1})$. Cette condition est encore équivalente à la condition (T) de Long [*8*] : pour tout processus adapté (γ_n) et tout $a > \gamma_0$, il existe un temps d'arrêt τ_a vérifiant les propriétés suivantes

$$\{\operatorname{Sup}_n \gamma_n > a\} \subset \{\tau_a < \infty\} \qquad \gamma_\tau \leq a \qquad \text{sur} \qquad \{\tau_a < \infty\}$$

et

$$P\{\tau_a < \infty\} \leq d.P\{\operatorname{Sup}_n \gamma_n > a\}$$

pour une constante $d \geq 1$.

Soit alors (X_n) une suite de ψ-médianes conditionnelles d'une v. a. $X \in M_+^\psi$, suivant une filtration $(\mathscr{F}_n)$ régulière et $\{\tau_a, a > X_0\}$, la famille des temps d'arrêt de Long associés au processus (X_n). On a visiblement $\int_{\{\tau<\infty\}} \psi(X - a) \leq 0$ pour tout $a > X_0$. D'où il résulte par les lemmes 8, 9, 15 et la propriété de régularité, que : $P\{\tau_a < \infty\} \leq dP\{X^* > a\}$, $a > X_0$; regroupons ces résultats :

THÉORÈME 19. — « *Inégalités maximales inverses* ». *On a toujours :*

$$\psi(a)\mathrm{P}\{\mathrm{X}^* > a\} \geq \frac{1}{2d}\int_{\{\mathrm{X}^*>a\}} \psi\left(\frac{\mathrm{X}}{2}\right)d\mathrm{P} \qquad \textit{pour} \quad a > \mathrm{X}_0 .$$

De plus si ψ *est modérée :*

$$\psi(a)\mathrm{P}\{\mathrm{X}^* > a\} \geq \frac{\beta}{\alpha d}\int_{\{\mathrm{X}^*>a\}} \psi(\mathrm{X})d\mathrm{P} \qquad \textit{etc...}$$

Remarque. — Lorsque $\psi(x) = x$, c'est-à-dire dans le cas de la théorie des martingales, on sait que toute filtration autorisant des inégalités maximales inverses est nécessairement régulière (cf. [*8*]) ; on ne peut donc affaiblir l'hypothèse de régularité faite ici. Terminons cette étude par un résultat analogue à celui de [*9*]. Soit ψ une fonction continue, croissante et impaire ; X une v. a. positive telle que $\psi(\mathrm{X}) \in \mathrm{L}^1$. Notons $(\psi(\mathrm{X}))_d$ une *réarrangée décroissante* de $\psi(\mathrm{X})$.

PROPOSITION 20. — « *Inégalités de Dubins et Gilat* ». *En notant par* $\mathrm{H}_\psi^\mathrm{X}$ *la fonction* $t \to \frac{1}{t}\int_0^t (\psi(\mathrm{X}))_d(y)dy$ *et par* λ *la mesure de Lebesgue. Si* $\mathrm{X}^* = \sup_n(\mathrm{X}_n)$, (X_n) *étant une suite de* ψ*-médiane conditionnelle de* X *relativement à la filtration croissante* $(\mathscr{F}_n)$ *; on a :*

$$\mathrm{P}\left\{\psi\left(\frac{\mathrm{X}^*}{2}\right) > 2y\right\} \leqq \lambda\{\mathrm{H}_\psi^\mathrm{X} > \mathrm{Y}\} \qquad \textit{pour tout} \quad y > 0 .$$

Démonstration. — Le théorème 18 montre que :

$$\frac{1}{2}\psi\left(\frac{a}{2}\right) \leq \frac{1}{\mathrm{P}\{\mathrm{X}^* > a\}}\int_{\{\mathrm{X}^*>a\}} \psi(\mathrm{X})d\mathrm{P} \leq \mathrm{H}_\psi^\mathrm{X}(\mathrm{P}\{\mathrm{X}^* > a\})$$

et comme la fonction $t \to \mathrm{H}_\psi^\mathrm{X}(t)$ est décroissante, il en résulte :

$$\mathrm{P}\{\mathrm{X}^* > a\} \leq \lambda\left\{\mathrm{H}_\psi > \frac{1}{2}\psi\left(\frac{a}{2}\right)\right\};$$

l'inégalité demandée s'en déduit.

Remarques. — 1°) On ne peut améliorer cette proposition. En effet si ψ est modérée, elle montre que $\mathrm{H}_\psi^\mathrm{X} \in \mathrm{L}^1(d\lambda) \Rightarrow \mathrm{X}^* \in \mathrm{L}^\psi$. Or, on a les équivalences $\mathrm{H}_\psi^\mathrm{X} \in \mathrm{L}^1(d\lambda) \Leftrightarrow \psi(\mathrm{X}) \in \mathrm{L}\log\mathrm{L}$ et $\psi(\mathrm{X}) \in \mathrm{L}\log\mathrm{L} \Leftrightarrow \mathrm{X}^* \in \mathrm{L}^\psi$, lorsque $(\mathscr{F}_n)$ est régulière. On retrouve ainsi un résultat de [*9*].

2°) On peut naturellement s'intéresser à l'espace $\mathrm{H}^\psi = \{\mathrm{X} ; \mathrm{X}^* \in \mathrm{L}^\psi\}$

d'après ce qui précède $H^{\psi} = \left\{ X\,; \text{il existe } \rho > 0 \text{ tel que } \psi\left(\frac{X}{\rho}\right) \in H^1 \right\}$.
On peut alors caractériser le dual de H^{ψ}, lorsque ψ est convexe modérée en utilisant la dualité de Young :

$$(H^{\psi})' = \{ X\,; \psi'(X) \in \mathrm{BMO} \}$$

où ψ' est la conjuguée de ψ.

BIBLIOGRAPHIE

[*1*] I. AMEMIYA and T. ANDO, Almost everywhere convergence of prediction sequences in $L_p(1 > p > \infty)$. *Z. f. W.*, t. **4**, 1965, p. 113-120.
[*2*] T. ANDO et T. SHINTANI, Best approximants in L_1-spaces. *Z. f. W.*, t. **33**, 1975, p. 33-39.
[*3*] D. BERNOULLI, Dijudicatio maxime probabilitis plurium observationum discrepantium atque verisimillima inductio inde formanda. *Acta Acad. Sci. Imp. Petrop. I*, 1777.
[*4*] G. BIRKHOFF, *Lattice theory*, AMS Colloquium publications, Vol. 25, 3e édition, 1979.
[*5*] D. BLACKWELL and L. E. DUBINS, A converse to the dominated convergence theorem. Ill. *J. Math.*, t. **7**, 1963, p. 508-514.
[*6*] B. BRU and H. HEINICH, Sur l'espérance des variables aléatoires vectorielles. *Ann. I. H. P.*, t. **16**, 1980, p. 177-220.
[*7*] B. BRU and H. HEINICH, Isométries positives et propriétés ergodiques de quelques espaces de Banach. *Ann. I. H. P.*, t. **4**, 1981, p. 377-405.
[*8*] B. BRU, H. HEINICH and J. C. LOOTGIETER, *Sur la régularité des filtrations*, *C. R. Acad. Sci.*, 1982, t. **294**, p. 313-316.
[*9*] L. E. DUBINS and D. GILAT, On the distribution of maxima of martingales. *Proc. Amer. Math. Soc.*, t. **68**, 1978, p. 237-238.
[*10*] M. H. DE GROOT and M. M. RAO, Multidimensional information inequalities. *Proc. Inter-Symp. Multiv. Anal. I.*, p. 287-313, Academic Press, 1967.
[*11*] H. HEINICH, Sur les mesures vectorielles signées. *C. R. Acad. Sci. Paris*, t. **289**, p. 285-286.
[*12*] N. HERRNDORF, Counterexamples to results of M. M. Rao, *Z. f. W.*, t. **53**, 1980, p. 283-290.
[*13*] N. HERRNDORF, Best Φ and N_{Φ}-approximants in Orlicz spaces of vector valued functions. *Z. f. W.*, t. **58**, 1981, p. 309-329.
[*14*] P. J. HUBER, *Théorie de l'inférence statistique robuste*. Presses Univ., Montréal, Département Math., 1969.
[*15*] D. LANDERS and L. ROGGE, Best approximants in L_{Φ}-spaces. *Z. f. W.*, t. **51**, 1980, p. 215-237.
[*16*] D. LANDERS and L. ROGGE, Isotonic approximation in L_s. *J. Approx. Theory*, t. **31**, 1981, p. 199-223.
[*17*] D. LANDERS and L. ROGGE, The natural median. *Ann. of Probab.*, t. **9**, 1981, p. 1041-1042.
[*18*] P. S. LAPLACE, *Mémoire sur la probabilité des causes par les événements*, 1774. Œuvres complètes VIII, p. 27-65.
[*19*] P. S. LAPLACE, Rechercher sur le milieu qu'il faut choisir entre les résultats de plusieurs observations, 1777. *Rev. Hist. Sci.*, t. **32**, 1979, p. 223-280.
[*20*] P. S. LAPLACE, *Mémoire sur les probabilités*, 1780. Œuvres complètes IX, p. 357-479.

[21] J. LINDENSTRAUSS and L. TZAFRIRI, *Classical Banach Spaces II*. Springer-Verlag, Berlin, Heidelberg, New York, 1979.
[22] J. L. LONG, Sur l'espace H_p de martingales régulières ($0 > p \leq 1$). *Ann. I. H. P.*, t. **17**, 1981, p. 123-142.
[23] J. NEVEU, *Bases mathématiques du Calcul des probabilités*. Masson, Paris, 1964.
[24] M. M. RAO, Inference in Stochastic Processes III. *Z.f. W.*, t. **8**, 1967, p. 49-72.
[25] M. M. RAO, Abstract non linear prediction and operator martingales. *J. Multiv. Anal.*, t. **1**, 1971, p. 129-157.

(Manuscrit reçu le 1er juillet 1983)
(Corrigé le 1er décembre 1984)