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

SAAB ABOU-JAOUDE

Conditions nécessaires et suffisantes de convergence L1 en probabilité de l'histogramme pour une densité

Annales de l'I. H. P., section B, tome 12, n° 3 (1976), p. 213-231

<http://www.numdam.org/item?id=AIHPB_1976__12_3_213_0>

Ann. Inst. Henri Poincaré,
Vol. XII, n° 3, 1976, p. 213-231.

Section B :
Calcul des Probabilités et Statistique.

Conditions nécessaires et suffisantes de convergence L1 en probabilité de l'histogramme pour une densité

par

SAAB ABOU-JAOUDE
2, rue de l'école des Postes 78000 Versailles

RÉSUMÉ. — Beaucoup de travaux ont eu pour objet la construction et l'étude d'estimateurs de la densité et notamment leur convergence simple et uniforme suivant un mode stochastique. Cet article traite de l'estimation d'une densité par la méthode de l'histogramme et cela, pour la distance L1. Des conditions nécessaires et suffisantes de convergence en probabilité sont données. Tous les résultats contenus dans cet article sont à notre connaissance originaux. L'étude du biais y est notamment abordée complètement.

SUMMARY. — The construction and study of density estimates has been the object of many scientific works, in particular, their simple and uniform convergence according to the stochastic type $\mathscr{M}$ has been well looked into. The following article deals with density estimate by the histogramm methode using the L1-distance. We give necessary and sufficient conditions for probability type convergence. The following results are to our knowledge original. The study of the bias has been dealt with thoroughly.

I. DÉFINITIONS ET NOTATIONS

— Soit $(\Omega, \mathscr{A}, \mu)$ un espace mesuré et $(\mathscr{P}_n)_{n\in\mathbb{N}^*}$ une suite de partitions de Ω. On note :

$$\mathscr{P}_n = \{ \Delta_{n,k} \mid k \in \mathbb{N}^* \}$$

(chaque partition $\mathscr{P}_n$ peut être finie ou dénombrable).

On suppose que les $\Delta_{n,k}$ vérifient la condition :

$$\forall n, k \in \mathbb{N}^*, \quad \mu(\Delta_{n,k}) < +\infty .$$

Notons que, sous cette condition, μ est σ-bornée.

— Soit f une densité sur $(\Omega, \mathscr{A}, \mu)$ c'est-à-dire une fonction $\mathscr{A}$-mesurable, positive, μ-intégrable avec $\mu(f) = 1$. On note P_f la probabilité associée à f.

On associe au couple $(f, \mathscr{P}_n)$ une fonction en escalier f_n définie de la manière suivante :

$$\forall k \in \mathbb{N}^*, \qquad \forall x \in \Delta_{n,k},$$

$$\begin{aligned} f_n(x) &= \frac{P_f(\Delta_{n,k})}{\mu(\Delta_{n,k})}, \qquad &\text{si} \quad \mu(\Delta_{n,k}) \neq 0 \\ &= 0, &\text{sinon} \end{aligned}$$

On a, évidemment :

$$f_n \geqslant 0 \qquad \text{et} \qquad \int_\Omega f_n d\mu = 1$$

DÉFINITION 1. — Nous dirons que la fonction f_n ainsi définie est la valeur moyenne de f sur la partition $\mathscr{P}_n$.

Cette définition généralise celle de valeur moyenne de f sur un élément de $\mathscr{A}$.

— Soit, d'autre part, $X_p = (x_1, x_2, \ldots, x_p)$ un échantillon de taille p de loi P_f. On associe à X_p et à la partition $\mathscr{P}_n$, une fonction $h_n(X_p)(\,.\,)$ définie par :

$$\forall k \in \mathbb{N}^*, \qquad \forall x \in \Delta_{n,k},$$

$$\begin{aligned} h_n(X_p)(x) &= \frac{1}{p\mu(\Delta_{n,k})} \operatorname{Card} \{ i \mid x_i \in \Delta_{n,k} \}, \qquad &\text{si} \quad \mu(\Delta_{n,k}) \neq 0 \\ &= 0, &\text{sinon} \end{aligned}$$

On a évidemment :

$$\int_\Omega h_n(X_p)(x) d\mu(x) \leqslant 1 ,$$

l'égalité ayant lieu P_f-presque sûrement.

DÉFINITION 2. — Nous dirons que $h_n(X_p)(.)$ est l'estimateur de la partition de f pour $\mathscr{P}_n$.

— Soit, enfin, $\mathscr{F}$ l'ensemble des densités sur $(\Omega, \mathscr{A}, \mu)$, d une distance sur $\mathscr{F}$, et $\mathscr{F}_0$ une partie de $\mathscr{F}_0$, munie de la distance induite. Pour $f \in \mathscr{F}_0$, on considère la v. a. r. (variable aléatoire réelle) :

$$\Delta(n, p) = d[h_n(X_p)(.), f(.)]$$

qui est l'écart entre f et son estimateur h_n sur $\mathscr{P}_n$. C'est une v. a. r. définie sur Ω^p.

— Enfin une définition commode, consacrée par l'usage ;

DÉFINITION 3. — Posons :

$$\Delta^{(1)}(n) \quad = d(f_n, f)$$

et

$$\Delta^{(2)}(n, p) = d[h_n(X_p)(.), f_n(.)]$$

La quantité $\Delta^{(1)}(n)$ est appelée le biais.

La quantité $\Delta^{(2)}(n, p)$ est appelée l'aléa.

Position du problème. — $\mathscr{M}$ étant un mode stochastique, le problème général de l'estimateur de la partition se pose de la manière suivante :

Quelles conditions doivent lier $\mathscr{F}_0$ et $(\mathscr{P}_n)$, et comment choisir n en fonction de p pour avoir :

$$\forall f \in \mathscr{F}_0, \quad \Delta[n(p), p] \underset{\mathscr{M}}{\to} 0 \qquad \text{pour } P_f. \quad (\alpha)$$

Plus particulièrement, on peut énoncer les deux questions suivantes :

— $\mathscr{F}_0$ étant donné, existe-t-il $(\mathscr{P}_n)$ telle que (α)? Ou, $(\mathscr{P}_n)$ étant donné, pour quelle famille $\mathscr{F}_0$ a-t-on (α) et sous quelles conditions sur $(\mathscr{P}_n)$?

— $\mathscr{F}_0$ et $(\mathscr{P}_n)$ étant choisis de telle sorte que le biais tende vers zéro, comment choisir n fonction de p pour avoir (α)?

Remarquons, tout d'abord, que la convergence du biais et de l'aléa vers zéro entraîne celle de $\Delta(n, p)$. Nous étudierons donc successivement :

— Des conditions de convergence du biais.

— Des conditions de convergence de l'aléa suivant le mode $\mathscr{M}$.

— La nécessité de la convergence du biais vers zéro.

Nous ferons cette étude dans le cadre suivant :

— d est la distance en variation entre probabilités associées aux densités ou, ce qui revient au même, la distance L1 sur $\mathscr{F}$: $d(f, g) = \int_\Omega | f-g | d\mu$.

— $\mathscr{M}$ est le mode en probabilité.

Remarque. — Le mode $\mathscr{M}$ ne concerne que l'aléa (lorsque le biais tend vers zéro) et il induit une condition sur le choix de n en fonction de p seulement.

Nous terminerons enfin cette étude par quelques applications à $\mathbb{R}$ et $\mathbb{R}^s$.

II. ÉTUDE DU BIAIS

Il s'agit ici d'étudier des conditions liant $\mathscr{F}_0$ et $\mathscr{P}_n$ pour que l'on ait :

$$\forall f \in \mathscr{F}_0, \qquad \lim_{n\to+\infty} \int_\Omega | f\text{-}f_n | d\mu = 0$$

Introduisons d'abord les notations suivantes :
Pour tout $n \in \mathbb{N}^*$,
— $\mathscr{B}_n$ est la σ-algèbre engendrée par $\mathscr{P}_n$,
— $\mathscr{B}'_n$ est la σ-algèbre engendrée par $\bigcup_{m\geqslant n} \mathscr{P}_m$,
— $\mathscr{B} = \bigcap_n \mathscr{B}'_n$;

Enfin, pour toute σ-algèbre $\mathscr{C}$, $\hat{\mathscr{C}}$ désigne la μ-complétée de $\mathscr{C}$.

Remarquons tout d'abord que, si le biais tend vers zéro (i. e. $f_n \xrightarrow[\mu]{L1} f$), il existe une suite extraite f_{n_j} convergeant vers f μ-presque partout. On en déduit que f est mesurable par rapport à $\hat{\mathscr{B}}$, cette condition n'étant évidemment pas suffisante. Pour établir une condition nécessaire plus forte, nous avons besoin de la définition suivante :

DÉFINITION 4. — On dit que $(\mathscr{P}_n)$ vérifie la propriété (C) d'approximation si :
$\forall A \in \mathscr{B}(0 < \mu(A) < +\infty)$, $\forall \varepsilon > 0$, $\exists n_0 \in \mathbb{N}^*$:

$$\forall n \geqslant n_0, \qquad \exists A_n \in \mathscr{B}_n : \mu(A \Delta A_n) \leqslant \varepsilon .$$

($A\Delta A_n$ = différence symétrique de A et A_n).

PROPOSITION 1. — Pour que, quelque soit f $\mathscr{B}$-mesurable, on ait : $f_n \xrightarrow[\mu]{L1} f$, il est nécessaire que $(\mathscr{P}_n)$ vérifie (C).

■ En effet, si $(\mathscr{P}_n)$ ne vérifiait pas (C), il existerait $A \in \mathscr{B}$ $(0 < \mu(A) < +\infty)$, $\varepsilon > 0$ et une suite n_j telle que :

$$\forall A_{n_j} \in \mathscr{B}, \qquad \mu(A \Delta A_{n_j}) > \varepsilon .$$

Nous pouvons supposer, pour alléger les notations, que la sous-suite n_j est la suite n tout entière.

Prenons, dans ces conditions, $f = \frac{1}{\mu(A)} \mathbb{1}_A = \lambda \mathbb{1}_A$ et posons :

$$\begin{aligned} J_n &= \{ k \mid \mu(\Delta_{n,k} \cap A) \neq 0, \quad \mu(\Delta_{n,k} \cap \bar{A}) \neq 0 \} \\ \alpha_{n,k} &= \mu(\Delta_{n,k} \cap A) \\ \beta_{n,k} &= \mu(\Delta_{n,k} \cap \bar{A}) \end{aligned}$$

Il est facile de voir que :

$$\begin{aligned} \int_\Omega | f_n - f | d\mu &= \sum_{k \in J_n} \int_{\Delta_{n,k}} | f_n - f | d\mu \\ &= \lambda \sum_{k \in J_n} \frac{2\alpha_{n,k}\beta_{n,k}}{\alpha_{n,k} + \beta_{n,k}} \end{aligned}$$

Posons :

$$J_n^{(1)} = \{ k \in J_n \mid \beta_{n,k} \leqslant \alpha_{n,k} \}$$

et

$$J_n^{(2)} = J_n - J_n^{(1)} .$$

Il vient :

$$\begin{aligned} \int_\Omega | f_n - f | d\mu &\geqslant \lambda \sum_{k \in J_n^{(1)}} \beta_{n,k} + \lambda \sum_{k \in J_n^{(2)}} \alpha_{n,k} \\ &\geqslant D_1 \qquad + D_2 \end{aligned}$$

On ne peut avoir à la fois $D_1 \leqslant \lambda\varepsilon/2$ et $D_2 \leqslant \lambda\varepsilon/2$ car il serait alors possible de construire $A_n \in \mathscr{B}_n$ tel que :

$$\mu(A_n \Delta A) \leqslant \varepsilon ,$$

(il suffit de prendre la réunion des $\Delta_{n,k} \subset A$ et des $\Delta_{n,k}$ avec $k \in J_n^{(1)}$.)

On en déduit que $\int_\Omega | f_n - f | d\mu$ ne tend pas vers zéro. ■

PROPOSITION 2. — La condition (C) est suffisante pour que, pour toute f $\mathscr{B}$-mesurable, le biais tende vers zéro.

■ La démonstration se fait en trois points.

1. Soit $A \in \mathscr{B}$ tel que $0 < \mu(A) < +\infty$. Démontrons la proposition pour $f = \mathbb{1}_A$.

On a, avec les notations de la proposition 1 :

$$\int_\Omega | f_n - f | d\mu = \sum_{k \in J_n} 2 \frac{\alpha_{n,k}\beta_{n,k}}{\alpha_{n,k} + \beta_{n,k}}$$

Par hypothèse, il existe une suite $A_n \in \mathscr{B}_n$ telle que :

$$\mu(A \Delta A_n) \to 0 \qquad \text{quand} \qquad n \to +\infty .$$

On peut supposer que, pour $k \in \mathrm{J}_n$,

$$\beta_{n,k} \leqslant \alpha_{n,k} \Rightarrow \Delta_{n,k} \subset \mathrm{A}_n$$

et

$$\beta_{n,k} > \alpha_{n,k} \Rightarrow \Delta_{n,k} \cap \mathrm{A}_n = \varnothing$$

Si ce n'était pas le cas, on modifierait A_n en y incluant ou en retranchant $\Delta_{n,k}$, ce qui aurait pour effet de diminuer $\mu(\mathrm{A}\Delta\mathrm{A}_n)$

Cela étant, on obtient :

$$\begin{aligned}\int_\Omega | f_n - f | d\mu &\leqslant 2 \sum_{k \in \mathrm{J}_n^{(1)}} \beta_{n,k} + 2 \sum_{k \in \mathrm{J}_n^{(2)}} \alpha_{n,k} \\ &\leqslant 2\mu(\mathrm{A}\Delta\mathrm{A}_n)\end{aligned}$$

Le résultat est alors une conséquence immédiate de cette inégalité.

2. On en déduit que, si f est en escalier sur $\mathscr{B}$ et ne prend qu'un nombre fini de valeurs $\left(f = \sum_{i=1}^{q} \lambda_i \mathbb{1}_{\mathrm{A}_i} \right)$, on a :

$$\begin{aligned}\int_\Omega | f_n - f | d\mu &= \int_\Omega | \mathrm{E}^{\mathscr{B}_n}(f) - f | d\mu \\ &= \int_\Omega \left| \sum_{i=1}^{q} \lambda_i (\mathrm{E}^{\mathscr{B}_n}(\mathbb{1}_{\mathrm{A}_i}) - \mathbb{1}_{\mathrm{A}_i}) \right| d\mu \\ &\leqslant \sum_{i=1}^{q} \lambda_i \int_\Omega | \mathrm{E}^{\mathscr{B}_n}(\mathbb{1}_{\mathrm{A}_i}) - \mathbb{1}_{\mathrm{A}_i} | d\mu \underset{n \to +\infty}{\longrightarrow} 0\end{aligned}$$

3. Soit, maintenant, f $\mathscr{B}$-mesurable positive. Il existe une suite g_p de fonctions en escalier sur $\mathscr{B}$, convergeant en croissant vers f et telle que g_p ne prenne qu'un nombre fini de valeurs. On sait qu'alors :

$$\int_\Omega | f - g_p | d\mu = \int_\Omega (f - g_p) d\mu \underset{p \to +\infty}{\longrightarrow} 0$$

De plus :

$$\mathrm{E}^{\mathscr{B}_n}(g_p) \leqslant \mathrm{E}^{\mathscr{B}_n}(f)$$

et

$$\int_\Omega \mathrm{E}^{\mathscr{B}_n}(f) d\mu = \int_\Omega f d\mu$$

$$\int_\Omega \mathrm{E}^{\mathscr{B}_n}(g_p) d\mu = \int_\Omega g_p d\mu .$$

On en déduit que $E^{\mathscr{B}_n}(g_p)$ converge L1 vers $E^{\mathscr{B}_n}(f)$ quand $p \to +\infty$, uniformément par rapport à n. D'après le théorème de limite commutative et le point 2, on a :

$$E^{\mathscr{B}_n}(f) = f_n \underset{\mu}{\overset{L1}{\rightarrow}} f. \quad \blacksquare$$

On peut donc énoncer le résultat suivant :

THÉORÈME 1. — Avec les notations ci-dessus, si $\mathscr{F}_0$ est le sous-ensemble de $\mathscr{F}$ de toutes les densités $\mathscr{B}$-mesurables, une C. N. S. pour que le biais tende vers zéro est que la suite $(\mathscr{P}_n)$ vérifie la propriété (C) d'approximation.

III. ÉTUDE DE L'ALÉA. CONVERGENCE EN PROBABILITÉ

Introduisons d'abord les notations suivantes :

Pour $M > 0$ et $n \in \mathbb{N}^*$, posons :

$$J_n(M) = \left\{ j \mid \mu(\Delta_{n,j}) \leqslant \frac{M}{n} \right\}$$

$$A(n, M) = \bigcup_{j \in J_n(M)} \Delta_{n,j}$$

et

$$a(n, M) = \mu(A(n, M))$$

$$a(M) = \limsup_n a(n, M)$$

et enfin,

$$a = \sup_M a(M) = \lim_{M \to +\infty} \uparrow a(M)$$

Supposons de plus que μ est finie ($\mu(\Omega) = 1$) et soit $\mathscr{F}_0$ un ensemble de densités sur $(\Omega, \mathscr{A}, \mu)$.

PROPOSITION 3. — Si $\mathbb{1}_\Omega \in \mathscr{F}_0$, une condition nécessaire de convergence de l'aléa P_f-en probabilité vers zéro est que :

$$\underline{a = 0}.$$

$\blacksquare$ Prenons en effet $f = \mathbb{1}_\Omega$ et supposons $a \neq 0$. Il existe alors $M > 0$ tel que $a(M) \geqslant 2\varepsilon$ et une suite n_j d'entiers $\uparrow +\infty$ telle que :

$$\mu(A(n_j, M)) \geqslant \varepsilon$$

Posons :

$$p_{n,k} = \mu(\Delta_{n,k})$$

et définissons la v. a. r. Z_n par :

$$Z_n = \sum_k p_{n,k} \lambda_{n,k}$$

ou $\lambda_{n,k}$ est la v. a. r. définie par :

$$\begin{aligned}\lambda_{n,k} &= 1 \quad \text{si } \Delta_{n,k} \text{ ne contient aucun point de l'échantillon } (x_1, x_2, \ldots, x_n) \text{ de loi } P_f\\ &= 0 \quad \text{sinon.}\end{aligned}$$

Z_n étant majoré par 1, la convergence de l'aléa vers zéro P_f-en probabilité entraîne celle de $E(Z_n)$.

Mais, on a :

$$\begin{aligned}E(Z_{n_j}) &= \sum_k p_{n_j,k} E(\lambda_{n_j,k})\\ &= \sum_k p_{n_j,k}(1 - p_{n_j,k})^{n_j}\\ &\geqslant \sum_{k\in J_{n_j}(M)} p_{n_j,k}\left(1 - \frac{M}{n_j}\right)^{n_j}\end{aligned}$$

Pour j suffisamment grand, on a :

$$E(Z_{n_j}) \geqslant \frac{e^{-M}}{2} \sum_{k\in J_{n_j}(M)} p_{n_j,k} = \frac{e^{-M}}{2} a(n_j, M)$$

Soit finalement :

$$E(Z_{n_j}) \geqslant \frac{e^{-M}}{2}\varepsilon .$$

L'aléa ne tend donc pas vers zéro, ce qui établit la nécessité de la condition $a = 0$. ■

PROPOSITION 4. — Toujours dans le cas $\mu(\Omega) = 1$, la condition $a = 0$ est suffisante pour entraîner la convergence P_f-en probabilité vers zéro de l'aléa et cela, pour toute densité f sur $(\Omega, \mathscr{A}, \mu)$.

■ Posons :

$$p_{n,k} = \int_{\Delta_{n,k}} f d\mu \quad \text{et} \quad \varphi_{n,k} = \int_{\Delta_{n,k}} h_n d\mu$$

On a :

$$\delta_n = d(f_n, h_n) = \sum_k | p_{n,k} - \varphi_{n,k} |$$

Si l'on pose :

$$\delta_n^+ = \sum_k (p_{n,k} - \varphi_{n,k})^+ ,$$

On a :

$$\delta_n = 2\delta_n^+$$

Il suffit, pour démontrer que l'aléa tend vers zéro P_f-en probabilité, d'établir que :

$$E_{P_f}(\delta_n^+) \to 0 \quad \text{quand} \quad n \to +\infty.$$

Ce résultat sera une conséquence du lemme suivant :

LEMME. — Soit X une v. a. $\mathscr{B}(n, p)$-variable de Bernouilli de paramètres n et p. Pour $n \geqslant 3$, $p \in]0,1[$ et $np > 1$, on a :

$$E\left(\left(p - \frac{X}{n}\right)^+\right) \leqslant \frac{1}{\sqrt{n\pi}}\sqrt{p} \tag{1}$$

et, évidemment, dans tous les cas :

$$E\left(\left(p - \frac{X}{n}\right)^+\right) \leqslant p \tag{2}$$

■ posant $s = [np]$ (partie entière de np), on a :

$$E((np - X)^+) = \sum_{r=0}^{s} (np - r)C_n^r p^r (1 - p)^{n-r}$$

$$= np\left[\sum_{r=1}^{s} (C_n^r p^r (1 - p)^{n-r} - C_{n-1}^{r-1} p^{r-1} (1 - p)^{n-r})\right] + np(1 - p)^n$$

$$= np(1 - p)^n + np\sum_{r=1}^{s} [C_{n-1}^r p^r (1 - p)^{n-r} - C_{n-1}^{r-1} p^{r-1}(1 - p)^{n-r+1}]$$

Soit en définitive :

$$\underline{E\left(\left(p - \frac{X}{n}\right)^+\right) = pC_{n-1}^s p^s (1 - p)^{n-s} \quad \text{avec} \quad s = [np].}$$

On déduit alors (1) de la formule ci-dessus en utilisant la formule de Stirling et en remarquant que :

$$\varphi(s) = \left(\frac{np}{s}\right)^s \left(\frac{nq}{n - s}\right)^{n-s} \quad \text{avec} \quad q = 1 - p$$

est maximum pour $s = np$ et que, par conséquent :

$\varphi(s) \leqslant 1$ ■

Utilisant le résultat de ce lemme, on obtient :

$$E(\delta_n^+) \leqslant \sum_{k \in J_n(M)} p_{n,k} + \frac{1}{\sqrt{n\pi}} \sum_{k \notin J_n(M)} \sqrt{p_{n,k}}\,.$$

En remarquant que $\psi = \text{Card}\,\overline{(J_n(M))} \leqslant \dfrac{n}{M}$ et que la fonction $x \to \sqrt{x}$ est concave, on obtient :

$$E(\delta_n^+) \leqslant P_f(A(n, M)) + \frac{1}{\sqrt{\pi M}}.$$

L'hypothèse entraîne que, pour M fixé, $\lim_{n\to+\infty} \mu(A(n, M)) = 0$. Le résultat se déduit alors de l'absolue continuité de P_f. ■

L'extension de ce résultat au cas μ σ-bornée nécessite l'introduction de notations différentes.

Pour $M > 0$, $n \in \mathbb{N}^*$ et $C \in \mathscr{A}$ de mesure finie, posons :

$$\begin{aligned}
\Delta_{n,k}(C) &= \Delta_{n,k} \cap C \\
J_n(M, C) &= \left\{ j \mid \mu(\Delta_{n,j}(C)) \leqslant \frac{M}{n} \right\} \\
A_n(M, C) &= \bigcup_{j \in J_n(M,C)} \Delta_{n,j}(C) \\
a_n(M, C) &= \mu(A_n(M, C)) \leqslant \mu(C) \\
a(M, C) &= \limsup_n a_n(M, C)
\end{aligned}$$

et, enfin :

$$a = \sup_{M,C} a(M, C)$$

DÉFINITION 5. — On dit que la suite $(\mathscr{P}_n)$ de partition de Ω vérifie la condition (D) si l'on a :

$$a = 0.$$

Nous pouvons maintenant énoncer une condition nécessaire de convergence de l'aléa.

PROPOSITION 3′. — Si, pour tout $C \in \mathscr{A}$ tel que $0 < \mu(C) < +\infty$, $\mathbb{1}_C/\mu(C)$ appartient à $\mathscr{F}_0$, la condition (D) est nécessaire pour avoir la convergence en probabilité de l'aléa vers zéro.

■ Supposons en effet $a \neq 0$ et soit (M, C) un couple tel que :

$$a(M, C) \geqslant 2\varepsilon.$$

Il existe une suite $n_j \uparrow +\infty$ telle que :

$$a_{n_j}(M, C) \geqslant \varepsilon$$

On a, de plus :

$$\mu(C) > 0.$$

Posons $\lambda = [\mu(C)]^{-1}$ et prenons :

$$f = \lambda \mathbb{1}_C \quad \text{et} \quad p_{n,k} = \lambda\mu(\Delta_{n,k}(C)).$$

Définissons alors la v. a. r. Z_n comme dans la proposition 3 en remplaçant $\Delta_{n,k}$ par $\Delta_{n,k}(C)$. Le raisonnement de la proposition 3 se répète mot pour mot à la seule différence que l'on arrive à l'inégalité :

$$\underline{E(Z_{n_j}) \geqslant \frac{\varepsilon}{2\lambda} e^{-M/\lambda}.}$$

La conclusion en découle. ■

Nous avons encore un résultat analogue à celui de la proposition 4.

PROPOSITION 4'. — La condition (D) est suffisante pour entraîner, si μ est σ-bornée, la convergence P_f-en probabilité vers zéro de l'aléa, et cela, pour toute densité f sur $(\Omega, \mathscr{A}, \mu)$.

■ $\varepsilon > 0$ étant fixé, le fait que μ soit σ-bornée entraîne l'existence d'un $C \in \mathscr{A}$ tel que :

$$0 < \mu(C) < +\infty \quad \text{et} \quad P_f(\bar{C}) \leqslant \varepsilon$$

C étant ainsi fixé, posons :

$$\Delta'_{n,k} = \Delta_{n,k} \cap C, \qquad p'_{n,k} = P_f(\Delta'_{n,k})$$
$$\Delta''_{n,k} = \Delta_{n,k} \cap \bar{C}, \qquad p''_{n,k} = P_f(\Delta''_{n,k})$$

et :

$$\varphi'_{n,k} = \int_{\Delta'_{n,k}} h_n d\mu, \qquad \varphi''_{n,k} = \int_{\Delta''_{n,k}} h_n d\mu$$

enfin :

$$\delta'_n = \sum_k | p'_{n,k} - \varphi'_{n,k} |, \qquad \delta''_n = \sum_k | p''_{n,k} - \varphi''_{n,k} |$$
$$\delta'^+_n = \sum_k (p'_{n,k} - \varphi'_{n,k})^+, \qquad \delta''^+_n = \sum_k (p''_{n,k} - \varphi''_{n,k})^+$$

On a, évidemment, pour tout n :

$$\delta^+_n \leqslant \delta'^+_n + \delta''^+_n$$

et, par construction de C,

$$\delta''^+_n \leqslant \varepsilon .$$

De plus, en reprenant la démonstration de la proposition 4,

$$E(\delta'^+_n) \leqslant P_f[A_n(M, C)] + \sqrt{\frac{\mu(C)}{\pi M}}$$
$$\left(\text{ici, card } \psi \leqslant \frac{n\mu(C)}{M}\right).$$

L'hypothèse entraîne que :

$$\lim_{n \to +\infty} \mu[A_n(M, C)] = 0 .$$

On conclut alors comme dans la proposition 4. ■

IV. NÉCESSITÉ DE LA CONVERGENCE DU BIAIS VERS ZÉRO

Nous allons établir le résultat plus général suivant dans le cas où μ est σ-bornée :

THÉORÈME 2. — Soit $(\mathscr{P}_n)$ une suite de partitions de Ω, f une densité sur $(\Omega, \mathscr{A}, \mu)$ et (g_n) une suite de fonctions à valeurs positives ou nulles et telle que g_n soit constante sur les éléments de $\mathscr{P}_n$ pour tout n.

Si (g_n) converge L1 vers f quand $n \to +\infty$, alors la suite f_n des valeurs moyennes de f sur $\mathscr{P}_n$ converge L1 vers f.

■ Soit $\mathscr{B}_n$ la σ-algèbre engendrée par $\mathscr{P}_n$.

1. Si g_n est une suite quelconque de fonctions convergeant L1 vers f, alors $E^{\mathscr{B}_p}(g_n)$ converge L1 vers $E^{\mathscr{B}_p}(f)$ uniformément par rapport à p. En effet :

$$| E^{\mathscr{B}_p}(g_n) - E^{\mathscr{B}_p}(f) | = | E^{\mathscr{B}_p}(f - g_n) | \leqslant E^{\mathscr{B}_p} | f - g_n |$$

D'où :

$$\int_\Omega | E^{\mathscr{B}_p}(g_n) - E^{\mathscr{B}_p}(f) | d\mu \leqslant \int_\Omega | g_n - f | d\mu$$

2. On a, de même :

$$\int_\Omega | E^{\mathscr{B}_p}(g_p) - E^{\mathscr{B}_p}(g_n) | \, d\mu \leqslant \int_\Omega | g_p - g_n | \, d\mu .$$

3. Majorons $\int_\Omega | f_p - f | \, d\mu$. On a :

$$\begin{aligned}\int_\Omega | f_p - f | \, d\mu &= \int_\Omega | E^{\mathscr{B}_p}(f) - f | \, d\mu \\ &\leqslant \int_\Omega | E^{\mathscr{B}_p}(f) - E^{\mathscr{B}_p}(g_n) | \, d\mu + \int_\Omega | E^{\mathscr{B}_p}(g_n) - E^{\mathscr{B}_p}(g_p) | \, d\mu \\ &\quad + \int_\Omega | E^{\mathscr{B}_p}(g_p) - f | \, d\mu .\end{aligned}$$

Finalement :

$$\int_\Omega |f_p - f| \, d\mu \leqslant 2 \int_\Omega |g_n - f| \, d\mu + 2 \int_\Omega |g_p - f| \, d\mu .$$

d'où le résultat. ■

On peut regrouper les résultats précédents dans l'énoncé suivant :

THÉORÈME 3. — Pour que l'estimateur de la partition de f pour $\mathscr{P}_n$ converge P_f-en probabilité pour la distance L1 quelle que soit la densité f sur $(\Omega, \mathscr{A}, \mu)$, il est nécessaire et suffisant que $\hat{\mathscr{B}} = \hat{\mathscr{A}}$ et que la suite $(\mathscr{P}_n)$ vérifie les propriétés (C) et (D).

Le problème de la compatibilité des conditions (C) et (D) est résolu dans la proposition suivante :

PROPOSITION 5. — On peut extraire d'une suite $(\mathscr{Q}_n)$ de partitions vérifiant la propriété (C) une suite $(\mathscr{P}_n)$ de partitions vérifiant à la fois (C) et (D).

■ Soit $(\lambda(n))$ une suite de réels > 0 tendant vers $+\infty$ et telle que $\frac{\lambda(n)}{n}$ soit décroissante et $(\eta(n))$ une suite de réels > 0 décroissant vers zéro.

Soit, de plus $\mathscr{Q}_n = \{ \Delta_{n,k} \mid k \in \mathbb{N}^* \}$ une suite de partitions vérifiant la propriété (C) et (C_n) une suite croissante d'éléments de $\mathscr{A}$ de mesure finie et tendant vers Ω.

Introduisons les notations suivantes :

$$J(n, p) = \left\{ j \mid \mu(\Delta_{p,j} \cap C_p) \leqslant \frac{\lambda(n)}{n} \right\}$$

$$A(n, p) = \bigcup_{j \in J(n,p)} \Delta_{p,j} \cap C_p$$

et $$a(n, p) = \mu(A(n, p)) .$$

Définissons, par récurrence sur n, la suite $\varphi(n)$ par $\varphi(1) = 1$, et, φ étant définie jusqu'au rang n,

— si $a(n+1, \varphi(n)+1) > \eta[\varphi(n)+1]$, $\quad \varphi(n+1) = \varphi(n)$

— si $a(n+1, \varphi(n)+1) \leqslant \eta[\varphi(n)+1]$, $\quad \varphi(n+1) = \varphi(n) + 1$

En remarquant que $a(n, p)$, à p fixé, tend en décroissant vers zéro, on déduit que : $\varphi(n) \uparrow +\infty$.

Posons alors :

$$\mathscr{P}_n = \mathscr{Q}_{\varphi(n)}, \quad D_n = C_{\varphi(n)} \quad \text{et} \quad \varepsilon(n) = \eta[\varphi(n)] .$$

$(\mathscr{P}_n)$ vérifie la propriété (C) et $D_n \uparrow \Omega$. Démontrons que la suite $(\mathscr{P}_n)$ ainsi construite vérifie la propriété (D) :

Remarquons d'abord, en utilisant les notations précédent la définition 5, que, par construction :

$$a_n(\lambda(n), D_n) \leqslant \varepsilon(n)$$

On en déduit aisément que, si C est un D_n et M un réel > 0,

$$\lim_{n \to +\infty} a_n(M, C) = 0$$

Soit maintenant $C \in \mathscr{A}$ de mesure finie et $\varepsilon > 0$. On a :

$$C = \lim \uparrow (D_n \cap C).$$

Il existe donc un entier n_0 tel que :

$$\mu(C - D_{n_0} \cap C) \leqslant \varepsilon.$$

Il vient alors :

$$a_n(M, C) \leqslant a_n(M, D_{n_0}) + \varepsilon.$$

On conclut en remarquant que :

$$\lim_{n \to +\infty} a_n(M, D_{n_0}) = 0 \qquad \blacksquare$$

V. APPLICATIONS

Dans ce paragraphe, nous utiliserons les définitions et notations introduites dans les paragraphes précédents.

Pour une partition $\mathscr{P}_n = (\Delta_{n,k})$ de Ω, posons :

$$\omega_n = \inf_k \mu(\Delta_{n,k})$$

DÉFINITION 6. — On dit qu'une suite $(\mathscr{P}_n)$ de partitions de Ω vérifie la condition (D') si l'on a :

$$\lim_{n \to +\infty} n\omega_n = +\infty$$

PROPOSITION 6. — Lorsque μ est finie, la condition (D') est suffisante pour entraîner la convergence de l'aléa P_f-en probabilité vers zéro, pour toute densité f sur Ω.

$\blacksquare$ En effet, la partition $\mathscr{P}_n$ est finie à partir d'un certain rang et, pour tout $M > 0$, $J_n(M)$ est vide à partir d'un certain rang. Le résultat découle alors de la proposition 4. $\blacksquare$

Remarque. — Lorsque μ n'est pas finie, la condition (D') ne suffit plus pour entraîner la convergence de l'aléa P_f-en probabilité vers zéro comme le montre l'exemple suivant :

— $\Omega = \mathbb{R}$, f est la densité uniforme sur [0,1], μ est la mesure de Lebesgue et $\Delta_{n,k}$ est défini par :

$$\Delta_{n,k} = \left[\frac{k-1}{n}, \frac{k}{n}\right[\cup [kn, (k+1)n[, \qquad k = 1, 2, \ldots, n$$

$$\Delta_{n,n+1} = \complement_{\mathbb{R}}\left(\bigcup_{k \leqslant n} \Delta_{n,k}\right)$$

Par contre, on a le résultat suivant :

PROPOSITION 6'. — Lorsque μ est σ-bornée relativement à $\mathscr{B}$, l'ensemble des deux conditions (C) et (D') suffit pour entraîner la convergence de l'aléa P_f-en probabilité vers zéro, pour toute densité f sur Ω.

(On dit que μ est σ-bornée relativement à $\mathscr{B}$, s'il existe une suite croissante d'éléments de $\mathscr{B}$ de mesure finie dont la réunion est égale à Ω.)

La démonstration est calquée sur celle de la proposition 4'.

■ $\varepsilon > 0$ étant fixé, il existe $C \in \mathscr{B}$ tel que :

$$\mu(C) < +\infty \quad \text{et} \quad P_f(\bar{C}) \leqslant \varepsilon .$$

La condition (C) entraîne l'existence d'une suite (A_n) telle que $A_n \in \mathscr{B}_n$ pour tout n et vérifiant :

$$\lim_{n \to +\infty} \mu(A_n \Delta C) = 0$$

Posons $I_n = \{ k \mid \Delta_{n,k} \subset A_n \}$ et, pour $k \in \bar{I}_n$, définissons $\Delta'_{n,k}$, $\Delta''_{n,k}$, $p'_{n,k}$, $p''_{n,k}$, $\varphi'_{n,k}$ et $\varphi''_{n,k}$ comme à la proposition 4', $p_{n,k}$ et $\varphi_{n,k}$ comme dans la proposition 4.

Posons :

$$\delta_n^1 = \sum_{k \in I_n} | p_{n,k} - \varphi_{n,k} |$$

$$\delta'_n = \sum_{k \in \bar{I}_n} | p'_{n,k} - \varphi'_{n,k} |$$

et

$$\delta''_n = \sum_{k \in \bar{I}_n} | p''_{n,k} - \varphi''_{n,k} |$$

On a :

$$\delta_n^+ \leqslant \delta_n^{1+} + \delta_n'^+ + \delta_n''^+$$

et,

$$E(\delta_n^+) \leqslant \sqrt{\frac{\mu(A_n)}{\pi n \omega_n}} + P_f(C - A_n) + \varepsilon$$

On conclut alors comme dans la proposition 4′ en utilisant (D′) et en remarquant que $\mu(A_n)$ est bornée. ■

De la proposition 6′, on déduit le théorème suivant :

THÉORÈME 4. — Supposons $\hat{\mathscr{B}} = \hat{\mathscr{A}}$ et la suite $(\mathscr{P}_n)$ vérifiant les conditions (C) et (D′). Alors, pour toute densité f sur $(\Omega, \mathscr{A}, \mu)$, l'estimateur de la partition converge P_f-en probabilité vers f, au sens de L1.

■ La proposition 6′ montre que l'aléa tend vers zéro, μ étant alors σ-bornée relativement à $\mathscr{A}$. La condition (C) assure, d'après la proposition 2, la convergence du biais vers zéro. ■

COROLLAIRE. — *Cas de $\mathbb{R}^s$ muni de la tribu $\mathscr{A}$ des boréliens et de la mesure μ de Lebesgue.* On considère dans $\mathbb{R}^s$ les partitions $Q_l = (\Delta_{l,q_1,q_2,\dots,q_s})$ définies par :

$$\Delta_{l,q_1,q_2,\dots,q_s} = \prod_{i=1}^{s} \left[\frac{q_i}{l}, \frac{q_{i+1}}{l}\right[, \qquad l \in \mathbb{N}^*, \quad (q_1, q_2, \dots, q_s) \in \mathbb{Z}^s,$$

et soit $l(n)$ une fonction de n. On considère la suite de partitions $\mathscr{P}_n = Q_{l(n)}$.

Pour que l'estimateur de la partition de f pour $(\mathscr{P}_n)$ converge P_f-en probabilité vers f, pour toute densité f sur $(\mathbb{R}^s, \mathscr{A}, \mu)$, il est suffisant que :

$$\lim_{n \to +\infty} \frac{n}{[l(n)]^s} = +\infty \qquad (D'')$$

■ En effet, (D″) n'est autre que la condition (D′). Il nous suffit donc d'établir que :

1. $\mathscr{B} \supset \mathscr{A}$
2. $\mathscr{A}$ vérifie la condition (C).

1. est évident. Il suffit de remarquer que $\mathscr{B}$ contient les pavés $\prod_{i=1}^{s}]-\infty, x_i[$ et que ces pavés engendrent $\mathscr{A}$.

Pour 2. il suffit d'établir que la propriété (C) d'approximation est vérifiée pour tout ouvert borné, la mesure μ étant régulière dans $\mathbb{R}^s$. (Tout borélien est limite décroissante d'ouverts.)

Soit donc $\mathscr{O}$ un ouvert borné de $\mathbb{R}^s$. $\mathscr{O}$ est union dénombrable de pavés ouverts de la forme $\prod_{i=1}^{s} [a_i, b_i[$. Pour tout $\varepsilon > 0$, il existe ainsi une famille finie (P_i), $i = 1, 2, \ldots, r$, de pavés ouverts tels que :

$$\mu\left(\mathcal{O} - \bigcup_{i=1}^{r} P_i\right) \leqslant \varepsilon .$$

Il nous suffit donc d'établir la propriété (C) pour une famille finie de pavés ouverts, donc pour un seul (propriété de la différence symétrique), et le résultat est élémentaire dans ce cas. ■

Voici, maintenant, un théorème d'existence d'estimateur convergeant L1 dans l'espace de toutes les densités sur Ω.

THÉORÈME 5. — Soit $(\Omega, \mathscr{A}, \mu)$ un espace mesuré, $\mathscr{F}$ l'ensemble de toutes les densités sur Ω. On suppose μ finie et $\mathscr{A}$ vérifiant les deux conditions suivantes :

H1 : Il existe une famille dénombrable $(A_i)_{i \in \mathbb{N}^*}$ d'éléments de $\mathscr{A}$ telle que :

$$\forall \varepsilon > 0, \forall A \in \mathscr{A}, \exists J \subset \mathbb{N}^* : \mu\left(A \Delta \left(\bigcup_{i \in J} A_i\right)\right) \leqslant \varepsilon .$$

H2 : $\mathscr{A}$ est engendrée par la famille (A_i).

Alors, il existe ~~un estimateur~~ convergeant P ~~-en probabilité vers~~ f ~~au~~ sens de L1, pour tout $f \in \mathscr{F}$. C'est l'estimateur de la partition pour une suite particulière $(\mathscr{P}_n)$ de partitions de Ω.

■ Remarquons d'abord que, d'après la proposition 5, il suffit de construire une suite de partitions vérifiant (C).

Posons pour cela :

$$Q_n = \{ A_i \mid i \leqslant n \} \cup \left\{ \complement_\Omega\left(\bigcup_{i=1}^{n} A_i\right) \right\},$$

et prenons pour $\mathscr{P}_n$ les atomes de Q_n. Par construction de $\mathscr{P}_n$, la suite $\mathscr{B}_n$ est croissante et l'on a, d'après H2 :

$$\mathscr{A} = \mathscr{B} = \bigcup_n \mathscr{B}_n .$$

Soit A un élément de $\mathscr{A}$. D'après H1, pour tout $\varepsilon > 0$, il existe $J \subset \mathbb{N}^*$ tel que si l'on pose :

$$B = \bigcup_{i \in J} A_i ,$$

on a :

$$\mu(A \Delta B) \leqslant \varepsilon .$$

Posons :

$$B_n = \bigcup_{\substack{i \in J \\ i \leqslant n}} A_i .$$

B_n appartient à $\mathscr{B}_n$ pour tout n. De plus, il existe un entier n_0 tel que :

$$\forall n \geqslant n_0 , \quad \mu(B - B_n) \leqslant \varepsilon$$

On en déduit :

$$\forall n \geqslant n_0 , \quad \mu(A \Delta B_n) \leqslant 2\varepsilon ,$$

C'est-à-dire la condition (C). ■

COROLLAIRE 1. — Ω est un espace topologique à base dénombrable, $\mathscr{A}$ la tribu de ses boréliens et μ une mesure finie sur Ω, régulière relativement à $\mathscr{A}$. Soit $\mathscr{F}$ l'ensemble des densités sur $(\Omega, \mathscr{A}, \mu)$ muni de la distance L1. Il existe alors dans $\mathscr{F}$ un estimateur convergeant P_f-en probabilité vers f, pour tout $f \in \mathscr{F}$.

■ $(\Omega, \mathscr{A}, \mu)$ vérifie alors les conditions du théorème 5. ■

Remarque. — On peut construire ici un estimateur de la partition convergeant P_f- presque complètement en remplaçant la condition (D') de la définition 6 par : $\lim_{n \to +\infty} \dfrac{n\omega_n}{\operatorname{Log} n} = +\infty$. Nous y reviendrons ultérieurement.

COROLLAIRE 2. — Sous les hypothèses du théorème 5, $\mathscr{F}$, muni de la distance L1, est séparable.

■ En effet, l'ensemble des fonctions en escalier, subordonnées à $\mathscr{P}_n$, à valeurs rationnelles est dénombrable et dense dans l'espace des fonctions en escalier subordonnées à $\mathscr{P}_n$ à valeurs réelles.

On conclut en remarquant que, pour tout $f \in \mathscr{F}$, il existe, d'après le théorème 5, une suite (f_n) de fonctions en escalier subordonnées à $(\mathscr{P}_n)$ et convergeant vers f au sens de L1. On utilise de plus le fait que $\mathscr{F}$ est métrique. ■

Étudions, enfin, le cas où Ω est dénombrable.

PROPOSITION 7. — On suppose que $\Omega = (a_r)_{r \in \mathbb{N}^*}$. Soit $P = (p_r)_{r \in \mathbb{N}^*}$ une probabilité sur Ω et X une v. a. à valeur dans Ω, de loi $P(Pr(X = a_r) = p_r)$. Soit $(X_1, X_2, \ldots, X_n)$ un n-échantillon de variable parente X et, pour tout $r \in \mathbb{N}^*$, v_r la fréquence de a_r dans l'échantillon.

Si l'on pose alors :

$$\delta_n = \sum_{r \in \mathbb{N}^*} \left| p_r - \frac{v_r}{n} \right| ,$$

δ_n converge P-en probabilité vers zéro, pour toute probabilité P sur Ω.