ANNALES DE LA FACULTÉ DES SCIENCES DE TOULOUSE

MARIANNE MORA

La convergence des fonctions variance des familles exponentielles naturelles

Annales de la faculté des sciences de Toulouse 5e série, tome 11, nº 2 (1990), p. 105-120

<http://www.numdam.org/item?id=AFST_1990_5_11_2_105_0>

Annales de la Faculté des Sciences de Toulouse Vol. XI, n° 2, 1990

La convergence des fonctions variance des familles exponentielles naturelles

MARIANNE MORA[1]

RÉSUMÉ. — Soit F une famille exponentielle naturelle sur un espace vectoriel E de dimension d. La famille F est caractérisée par la donnée du domaine J_F des moyennes m des probabilités $P(m,F)$ qui la composent et par la fonction variance V_F qui à tout m de J_F fait correspondre $V_F(m)$ l'opérateur de variance de la loi $P(m,F)$. Nous montrerons que, sous certaines conditions, la fonction limite d'une suite de fonctions variance est encore la fonction variance d'une famille exponentielle naturelle.

ABSTRACT. — Let F be a natural exponential family on a d-dimensional vector space E. Denoting by J_F the domain of the means of the probability measure $P(m,F)$ in F with mean m, let $V_F(m)$ be the variance operator of the probability measure $P(m,F)$. Then the function $V_F : m \in J_F \mapsto V_F(m)$ is called the variance function of F; under fairly large conditions, it is shown that the limit function of a sequence of variance functions is still the variance function of a natural exponential family.

KEY-WORDS : Natural Exponential Families
Variance Functions
Limits of Laplace Transforms

AMS – Classification 60B10

0. Introduction

L'article traite de familles exponentielles naturelles définies sur un espace vectoriel réel, de dimension finie. Il est bien connu, et facile à vérifier, qu'une telle famille F est caractérisée par sa fonction variance V_F. En pratique ces fonctions variance sont souvent très simples (voir par exemple l'article

[1] Marianne Mora, U.F.R. Sciences Économiques, Université de Paris X Nanterre, 200, avenue de la République, 92001 Nanterre Cedex

de Morris [6], ou [3] et [4]) et s'avèrent aussi maniables pour identifier une famille que l'est, par exemple, une transformée de Fourier pour identifier une loi de probabilité. Il serait donc intéressant de disposer, pour ces variances, d'une sorte de "théorème de Paul Lévy" disant quand la limite d'une suite de fonctions variance est encore une fonction variance : ce problème est d'ailleurs esquissé dans quelques lignes de l'article de Morris [6]. Ce théorème de convergence (section 3) n'est qu'un outil technique, et dans le présent article, nous n'en présenterons pas d'applications; mais nous pensons à la classification des familles exponentielles par leurs fonctions variance : par exemple aux variances quadratiques à plusieurs dimensions [5] ou aux variances sur $\mathbb{R}$ de la forme $P + Q\sqrt{R}$ avec P, Q, R polynômes de degré inférieur ou égal à 3, 2 et 2 respectivement.

Par exemple, on peut vérifier facilement que $V(m) = \sqrt{m}$ sur $]0, +\infty[$ n'est pas une fonction variance. Notre résultat impliquera alors que $\sqrt{m}\,(1+\epsilon P(m))$ pour P polynôme et ϵ petit ne peut être une fonction variance, chose difficile à voir directement.

Ce problème de "passage à la limite" s'est posé de façon naturelle dans un travail précédent [4] qui concernait l'étude et la classification de toutes les fonctions variance polynomiales de degré 3. Dans ce problème de classification, il est facile de prouver, à l'aide d'une formule de Lagrange, que la fonction $V_F(m) = m(m+p)(m+p_1)$, définie sur $J =]0, +\infty[$ et pour p et $p_1 > 0$, est une fonction variance. Il est alors tentant d'en déduire, par passage à la limite, en faisant tendre p_1 vers 0, que la fonction $V_F(m) = m^2(m+p)$ est aussi une fonction variance. Ce résultat a été prouvé de trois façon différentes : tout d'abord par une démonstration due à Paul Ressel et décrite dans [4], puis par un calcul explicite (G. Letac, [2]), à l'aide d'un certain processus stochastique et enfin grâce à un théorème de Jorgensen et d'un corollaire dû à Barlev décrits dans [3]. Nous en donnons ici une quatrième grâce au théorème de convergence pour les fonctions variance. La section 1 comprend des rappels sur les familles exponentielles naturelles. Dans la section 2 (plutôt qu'en appendice) certains points d'analyse réelle sont énoncés, enfin la section 3 est consacrée à l'énoncé et à la démonstration du théorème de convergence.

1. Familles exponentielles naturelles et fonctions variance sur un espace vectoriel de dimension finie

Soit E un espace vectoriel de dimension finie d, soit E^* son espace dual. Si μ est une mesure σ-finie, positive, sur E, on définit pour tout $\theta \in E^*$, la fonction :

$$k_\mu(\theta) = \log \int \exp(\langle \theta, x \rangle) \mu(\mathrm{d}x) .$$

Nous noterons :

- Θ_μ l'intérieur de l'ensemble convexe $\{\theta \in E^* \mid k_\mu(\theta) < \infty\}$,
- C_μ le plus petit convexe fermé de E tel que son complémentaire soit de μ-mesure nulle,
- I_μ l'intérieur de C_μ.

Enfin nous considérons l'ensemble :

$$M(E) = \{\mu \text{ mesure } \sigma\text{-finie, positive sur } E \text{ tel que } \Theta_\mu \neq \emptyset \text{ et } I_\mu \neq \emptyset\} .$$

DÉFINITION 1.1.— *Pour μ dans $M(E)$, la famille de lois de probabilité sur E, $F(\mu)$, définie par :*

$$F(\mu) = \{P(\theta, \mu) \,,\ \theta \in \Theta_\mu\}$$

où $P(\theta,\mu)(\mathrm{d}x) = \exp\{\langle \theta, x\rangle - k_\mu(\theta)\}\mu(\mathrm{d}x)$, $\forall\, \theta \in \Theta_\mu$, est appelée famille exponentielle naturelle engendrée par μ.

Remarque.— Il est facile de voir que pour μ_1 et μ_2 dans $M(E)$, on a l'équivalence suivante :

$$F(\mu_1) = F(\mu_2) \Leftrightarrow$$
$$\Leftrightarrow \exists\ (a, b) \in E^* \times \mathbb{R} \text{ tel que } \mu_1(\mathrm{d}x) = \exp(\langle a, x\rangle + b)\, \mu_2(\mathrm{d}x) .$$

DÉFINITION 1.2.— *Une famille F de lois de probabilité sur E est appelée famille exponentielle naturelle sur E s'il existe une mesure μ dans $M(E)$ telle que $F = F(\mu)$.*

Nous omettrons dans ce qui suit le terme "naturel". La paramétrisation de la famille $F = F(\mu)$ par l'ouvert Θ_μ est donc liée à un choix particulier de mesure μ. Cependant il est possible de définir une autre paramétrisation de F à l'aide des moyennes. Pour cela il est nécessaire de rappeler certaines propriétés de la fonction k_μ, dont nous omettons les démonstrations, très faciles (voir [3]).

PROPOSITION 1.1.— *Soit $\mu \in M(E)$, alors :*

(1) Θ_μ est convexe;

(2) k_μ est strictement convexe, analytique-réelle sur Θ_μ ;

(3) $\forall\, \theta \in \Theta_\mu$, la dérivée de k_μ en θ définie par :

$$k'_\mu(\theta) = \int xP(\theta,\mu)(\mathrm{d}x)$$

appartient à I_μ et la fonction dérivée k'_μ est bijective de Θ_μ dans $J_\mu = k'_\mu(\Theta_\mu)$;

(4) $\forall\, \theta \in \Theta_\mu$, $\forall\, (h_1,h_2) \in (E^*)^2$ et si $m = k'_\mu(\theta)$, alors :

$$k''_\mu(\theta)\cdot(h_1,h_2) = \int h_1(x-m)\,h_2(x-m)P(\theta,\mu)(\mathrm{d}x)\,.$$

De plus l'application : $h \mapsto k''_\mu(\theta)\cdot(h,h)$ est une forme quadratique strictement définie positive sur E^*.

Notons alors ψ_μ l'application réciproque de k'_μ définie sur J_μ. Il est possible de donner une nouvelle définition de la famille $F(\mu)$, sous la forme :

$$F(\mu) = \left\{P(\psi_\mu(m)\,,\,\mu),\quad m \in J_\mu\right\}\,.$$

Nous appellerons J_μ le domaine des moyennes de $F(\mu)$. C'est une partie ouverte et connexe de E.

PROPOSITION 1.2.— *Soient μ_1 et μ_2 dans $M(E)$. Alors :*

$$\begin{aligned} F(\mu_1) = F(\mu_2) \quad &\Rightarrow \textit{1) } I_{\mu_1} = I_{\mu_2} = I_F \\ &\Rightarrow \textit{2) } J_{\mu_1} = J_{\mu_2} = J_F \\ &\quad\;\; \textit{3) } \forall\, m \in J_F : P\big(\psi_{\mu_1}(m)\,,\,\mu_1\big) = P\big(\psi_{\mu_2}(m)\,,\,\mu_2\big)\,. \end{aligned}$$

On a alors la définition suivante :

DÉFINITION 1.3.— *Soit F une famille exponentielle sur E et μ dans $M(E)$ telle que $F = F(\mu)$.*

L'application $m \mapsto P(m,F)$, de J_F dans F, définie par :

$$P(m,F)(\mathrm{d}x) = P\big(\psi_\mu(m)\,,\,\mu\big)(\mathrm{d}x)$$

est appelée paramétrisation canonique de F par son domaine des moyennes.

On peut alors introduire la fonction variance d'une famille exponentielle F.

DÉFINITION 1.4. — *Soit $F = \{P(m,F)\,,\ m \in J_F\}$ une famille exponentielle sur E. L'application V_F définie par :*

$$\begin{aligned} V_F : J_F &\mapsto Q(E^*) = \{\textit{formes quadratiques sur } E^*\} \\ m &\mapsto V_F(m) = \textit{variance de } P(m,F) \end{aligned}$$

est appelée fonction variance de F.

Cette application a les propriétés suivantes :

PROPOSITION 1.3. — *Soit F une famille exponentielle sur E. Soit V_F sa fonction variance définie sur J_F. Alors :*

1) $V_F(m) = k''_\mu(\psi_\mu(m)) = \left(\psi'_\mu(m)\right)^{-1}$, $\forall\, m \in J_F$, $\forall\, \mu \in M(E)$ telle que $F = F(\mu)$.

2) De plus si F et F' sont deux familles exponentielles sur E :

$$F = F' \Leftrightarrow J_F = J_{F'} \quad \textit{et} \quad V_F = V_{F'}\,.$$

Cette proposition montre que (J_F, V_F) caractérise la famille F.

2. Deux résultats d'analyse réelle

Dans la démonstration du théorème de convergence (section 3), nous aurons besoin des deux points suivants, concernant la différentiabilité de fonctions définies sur un espace de Banach E ; ils sont énoncés et démontrés dans le livre de Cartan ([1] théorème 3.6.1, p. 49 et théorème 3.4.1 p. 48) :

THÉORÈME 2.1

Soit U un ouvert convexe d'un espace de Banach E et soit $\left(f_n\right)_{n\geq 1}$ une suite d'applications différentiables sur U, et à valeurs dans un espace de Banach F. On suppose que :

1) La suite des applications dérivées $\left(f'_n\right)_{n\geq 0}$

$$f'_n : U \to \mathcal{L}(E,F)$$

converge, uniformément dans U, vers une application g

$$g : U \to \mathcal{L}(E,F)\,.$$

2) Il existe $a \in U$ tel que $f(a) = \lim_{n\to\infty} f_n(a)$ existe. Alors :

$$\forall\, x \in U\,, \quad f(x) = \lim_{n\to\infty} f_n(x) \text{ existe}$$

et la convergence de la suite $\left(f_n\right)_{n\geq 0}$ vers f est uniforme sur chaque partie bornée de U.

De plus f est différentiable sur U avec $f' = g$.

THÉORÈME 2.2

Soit U un ouvert connexe d'un espace de Banach E. Pour tout $(a,b) \in U^2$ on définit la distance $d_U(a,b)$ par :

$$d_U(a,b) = \inf \{\text{longueurs des lignes brisées d'extrémités } a \text{ et } b \text{ incluses dans } U\}.$$

Soit $f : U \to F$ (Banach), différentiable et telle qu'il existe $k > 0$:

$$\|f'(x)\| \leq k\,.$$

Alors : $$\forall\, (x_1,x_2) \in U^2\,, \quad \|f(x_1) - f(x_2)\| \leq k\, d_U(x_1,x_2)\,.$$

3. Théorème de convergence pour des suites de fonctions variance

Voici le résultat de l'article :

THÉORÈME 3.1

Soit $\left(F_n\right)_{n\geq 1}$ une suite de familles exponentielles sur E, de domaine des moyennes et de fonctions variance respectifs J_n et V_n.

On suppose que :

1) Il existe un ouvert convexe de E non vide J tel que :

$$J \subset \bigcap_{n\geq 1} J_n\,.$$

2) Pour tout m dans J, $V(m) = \lim_{n\to\infty} V_n(m)$ existe, $V(m)$ est une forme quadratique sur E^* strictement positive et la suite $\left(V_n\right)_{n\geq 1}$ converge vers V, uniformément sur tout compact contenu dans J.

Alors :

i) il existe une famille exponentielle F sur E telle que :

$$V_F(m) = V(m)\,, \quad \forall\, m \in J\,;$$

ii) pour tout m dans J, la suite $\big(P(m, F_n)\big)_{n\geq 1}$ converge vers $P(m, F)$, au sens de la convergence étroite.

Démonstration. — Cette démonstration comporte deux parties. Dans la partie (A) nous montrons que pour tout $m \in J$ et, si l'on note :

$$\mu_n(m) = P(m, F_n) \quad \text{et} \quad k_n = k_{\mu_n}(m)\,,$$

alors la suite $\big(k_n\big)_{n\geq 1}$ converge vers la fonction $k_{\mu(m)}$, fonction génératrice des cumulants d'une mesure $\mu(m)$ dans $M(E)$. De plus la mesure $\mu(m)$ est telle que sa moyenne et sa variance sont respectivement m et $V(m)$.

Dans la seconde partie (B) de la démonstration nous montrerons qu'il existe une famille exponentielle F telle que $F = F(\mu(m))$, $\forall\, m \in J$.

A) 1) $\forall\, m \in J$, $\forall\, n \geq 1$, $V_n(m)$ et $V(m)$ peuvent être considérés comme des isomorphismes de E^* dans E, étant des formes quadratiques strictement positives sur E^*. Nous noterons respectivement $\psi'_n(m)$ et $\psi'(m)$ leurs isomorphismes réciproques, bien que nous n'ayons pas encore montré que ψ' est une différentielle. Nous pouvons ainsi définir les fonctions ψ'_n et ψ' correspondantes en posant :

$$\begin{cases} \psi'_n : J \to \operatorname{Isom}(E, E^*) & n \geq 1 \\ m \mapsto \psi'_n(m) \end{cases} \qquad \text{et} \qquad \begin{cases} \psi' : J \to \operatorname{Isom}(E, E^*) \\ m \mapsto \psi'(m)\,. \end{cases}$$

On a alors le résultat suivant :

Pour un compact K fixé dans J, la suite $\big(\psi'_n\big)_{n\geq 1}$ converge vers ψ', uniformément sur K. (1)

En effet si E et F sont des espaces de Banach, et si un opérateur $u : E \to \mathcal{L}(E, F)$ est inversible, alors :

$$\left\|u^{-1}\right\|^{-1} = \inf\{c \geq 0 : \|u(x)\| \geq c\|x\|\,,\ \forall\, x \in E\}$$

où :

$$\|u\| = \sup_{\substack{x\in E \\ \|x\|=1}} \|u(x)\|\,. \tag{2}$$

Donc, puisque :

$$\begin{aligned}\left\|\psi_n'(m)-\psi'(m)\right\| &= \left\|\left(V_n(m)\right)^{-1}-\left(V(m)\right)^{-1}\right\| \\ &\leq \left\|\left(V_n(m)\right)^{-1}\right\| \left\|\left(V(m)\right)^{-1}\right\| \left\|V_n(m)-V(m)\right\| ,\end{aligned}$$

pour prouver (1), il suffit de prouver que la suite $\left\|\left(V_n(m)\right)^{-1}\right\|$ est uniformément bornée sur K.

Cela est conséquence immédiate de l'hypothèse (2) du théorème et de la formule (2) relative aux opérateurs inversibles.

Ainsi, la suite $\left(\psi_n'\right)_{n\geq 1}$ converge vers ψ', uniformément sur K.

2) Fixons maintenant m_0 dans J et considérons un ouvert convexe borné J_1 vérifiant :

$$m_0 \in J_1 \quad \text{et} \quad J_1 \subset \overline{J_1} \subset J .$$

Cet ouvert J_1 dépend donc du point m_0 choisi dans J, ($J_1 = J_1(m_0)$).

Soit :

$$\mu_n = P(m_0, F_n) \quad \text{et} \quad \psi_n = \psi_{\mu_n} , \quad \forall\, n \geq 1 .$$

Alors :

a) $\forall\, m \in J_1, \quad f(m) = \lim_{n\to\infty} \psi_n(m)$ existe.

b) $\left(\psi_n\right)_{n\geq 1}$ converge vers f, uniformément sur tout compact de J_1. De plus f est différentiable sur J_1 avec :

$$f' = \psi' .$$

En effet, puisque $\forall\, m \in J_1$ et $\forall\, n \geq 1$ on a par construction :

$$\psi_n(m_0) = 0 \quad \text{et} \quad \psi_n'(m) = \left(V_n(m)\right)^{-1} ,$$

la suite $\left(\psi_n\right)_{n\geq 1}$ vérifie les hypothèses du théorème 1 de la section 2 et l'on conclut immédiatement.

Nous noterons désormais ψ la fonction limite f ; on a alors $\psi(m_0) = 0$.

3) Soit $I_n = \psi_n(J_1)$, $n \geq 1$.

Puisque I_n est ouvert, borné et contient 0 pour tout $n \geq 1$, on peut définir pour tout $n \geq 1$:

$$r_n = \inf\{\|x\| ,\, x \in F_r(I_n)\}$$

où $F_r(I_n)$ désigne la frontière de I_n.

Alors, $\forall\, n \geq 1$, $B(0, r_n) \subset I_n$.

Montrons alors que $\inf_{n\to\infty} r_n > 0$. Sinon $\underline{\lim}\, r_n = 0$ et sans perte de généralité on peut alors supposer que $\lim r_n = 0$.

Mais, puisque $\forall\, n \geq 1$, ψ_n est bijective, bicontinue sur J, pour tout ouvert A on a :

$$\psi_n(\overline{A}) = \overline{\psi_n(A)}$$

et donc :

$$\psi_n\left(F_r(J_1)\right) = \psi_n\left(\overline{J_1} \cap \overline{\complement J_1}\right) = \overline{\psi_n(J_1)} \cap \overline{\complement \psi_n}(J_1) = F_r(I_n)\,.$$

Puisque $x \mapsto \|x\|$ est continue sur E, pour tout $n \geq 1$, il existe x_n dans $F_r(I_n)$ et donc m_n dans $F_r(J_1)$ tels que :

$$\|\psi_n(m_n)\| = r_n\,.$$

$F_r(J_1)$ étant compact, on peut alors supposer sans perte de généralité qu'il existe m dans $F_r(J_1)$ tel que :

$$\lim_{n\to\infty} \|m_n - m\| = 0\,.$$

Ainsi :

$$\begin{aligned}\|\psi(m)\| &\leq \|\psi(m) - \psi_n(m)\| + \|\psi_n(m) - \psi_n(m_n)\| + r_n \\ &\leq \|m_n - m\|\; \|\psi_n'(\alpha_n)\| + \|\psi(m) - \psi_n(m)\| + r_n\end{aligned}$$

pour un α_n dans $[\,m\,,\,m_n\,]$.

Enfin, puisque $\|\psi_n'(\alpha_n)\|$ est uniformément bornée sur J_1, et que :

$$\lim_{n\to\infty} r_n = \lim_{n\to\infty} \|m_n - m\| = \lim_{n\to\infty} \|\psi_n(m) - \psi(m)\| = 0\,,$$

on conclut que $\|\psi(m)\| = 0$, ce qui est impossible car ψ est bijective et $m \neq m_0$.

Dans la suite nous fixerons r tel que $0 < r < \inf r_n$. On aura donc :

$$B(0, r) \subset \bigcap_{n\geq 1} \psi_n(J_1)\,.$$

4) Considérons pour tout $n \geq 1$, les fonctions $k_n' = k_{\mu_n}'$ définies respectivement sur $\psi_n(J_1)$.

De la même façon on peut définir k' comme la fonction réciproque de ψ (bien que nous n'ayons pas encore montré que k' est une différentielle).

Montrons que $\left(k'_n\right)_{n\geq 1}$ *converge vers* k'*, uniformément sur* $B\left(0\,,\,\dfrac{r}{2}\right)$ *:*

Soit $\theta \in B\left(0\,,\,{}^{r}/_{2}\right)$. Soit $m = k'(\theta)$ dans J, et soit pour tout $n \geq 1$, $\theta_n = \psi_n(m)$ dans $\psi_n(J_1)$. Alors :

$$k'_n(\theta) - k'(\theta) = k'_n(\theta) - k'_n(\theta_n)$$

et

$$\left\|k''_n(\theta)\right\| = \|V_n(m_n)\|\,, \quad \text{où } m_n = k'_n(\theta) \in J_1\,, \quad \forall\, n \geq 1\,.$$

Puisque V_n est uniformément bornée sur J_1, il existe $M > 0$ tel que :

$$\left\|k''_n(\theta)\right\| \leq M\,, \quad \forall\,\theta \in B\left(0\,,\,\frac{r}{2}\right)\,,\ \forall\, n \geq 1\,.$$

J_n étant connexe, on peut appliquer le théorème 2.2 et on a :

$$\left\|k'_n(\theta) - k'_n(\theta_n)\right\| \leq M d_n\left(\psi_n(m), \psi(m)\right)$$

où $d_n(\,\cdot\,,\,\cdot\,) = d_{I_n}(\,\cdot\,,\,\cdot\,)$.

De plus $\left(\psi_n\right)_{n\geq 1}$ converge vers ψ, c'est-à-dire :

$$\forall\,\epsilon \leq \frac{r}{2}\,,\ \exists\, N\left(\epsilon, \overline{J_1}\right) \quad : \quad \forall\, n \geq N\,,\ \|\psi_n(m) - \psi(m)\| \leq \epsilon\,,\ \forall\, m \in J_1\,.$$

Donc $\forall\, n \geq N$, $\psi_n(m) \in B\left(\psi(m)\,,\,{}^{r}/_{2}\right)$ et ainsi :

$$\forall\, n \geq N\,, \quad d_n\left(\psi_n(m)\,,\,\psi(m)\right) = \|\psi_n(m) - \psi(m)\| \leq \epsilon\,, \quad \forall\, m \in J_1\,.$$

En particulier pour $\epsilon = {}^{r}/_{2}$, il existe $N\left(({}^{r}/_{2})\,,\,J_1\right)$ tel que :

$$\forall\, n \geq N \quad : \quad d_n\left(\psi_n(m)\,,\,\psi(m)\right) \leq \epsilon\,.$$

Supposons maintenant qu'il existe $\epsilon_0 \geq {}^{r}/_{2}$ tel que :

$$\forall\, N \geq 1\,,\ \exists\, n \geq N\,,\ \exists\, m_n \in J_1 \quad : \quad d_n\left(\psi_n(m_n)\,,\,\psi(m_n)\right) \geq \epsilon_0\,. \qquad (*)$$

Donc, pour $N = N({}^{r}/_{2})$, il existe $n_0 > N$ et m_0 dans J_1 tels que :

$$d_{n_0}\left(\psi(m_0)\,,\,\psi_{n_0}(m_0)\right) \geq \epsilon_0$$

et

$$d_{n_0}\left(\psi(m_0)\,,\,\psi_{n_0}(m_0)\right) = \|\psi_{n_0}(m_0) - \psi(m_0)\| \geq \frac{r}{2}\,.$$

Cela implique que $\|\psi_{n_0}(m_0) - \psi(m_0)\| = {}^{r}/_{2}$.

Ainsi, sous l'hypothèse $(*)$ on peut construire une suite (strictement) croissante d'entiers $\left(n_k\right)_{k\geq 1}$ et une suite de points $\left(m_k\right)_{k=1}^{\infty}$ dans J, telles que :

$$\forall\, k \geq 1\,, \quad d_{n_k}\left(\psi_{n_k}(m_k)\,,\,\psi(m_k)\right) = \frac{r}{2} = \|\psi_{n_k}(m_k) - \psi(m_k)\|\,.$$

Soit m dans $\overline{J}_1$ telle que $\left(m_k\right)_{k\geq 1}$ converge vers m. Alors :

$$\begin{aligned}&\|\psi_{n_k}(m_k) - \psi(m_k)\| \\ &\quad \leq \|m_k - m\| \left(\left\|\psi_n'(\alpha_k)\right\| + \left\|\psi'(\beta_k)\right\| + \|\psi_{n_k}(m) - \psi(m)\|\right)\end{aligned}$$

pour α_k et β_k dans $\overline{J}_1$.

Enfin, de la convergence uniforme de $\left(\psi_{n_k}\right)_{k\geq 1}$ vers ψ sur $\overline{J}_1$, de la convergence de $\left(m_k\right)_{k\geq 1}$ vers m, et sachant que ψ_{n_k}' et ψ' sont uniformément bornées sur $\overline{J}_1$, on déduit que :

$$\forall\, \epsilon > 0\,,\ \exists\, N(\epsilon) \quad : \quad \forall\, n \geq N\,,\ \|\psi_{n_k}(m_k) - \psi(m_k)\| = \frac{r}{2} \leq \epsilon\,,$$

ce qui est impossible. L'hypothèse $(*)$ est donc fausse, c'est-à-dire que :

$$\forall\, \epsilon > 0\,,\ \exists\, n \geq N \quad : \quad d_n\left(\psi_n(m)\,,\,\psi(m)\right) \leq \epsilon\,,\ \forall\, m \in J_1\,.$$

On a ainsi prouvé la convergence uniforme de $\left(k_n'\right)_{n\geq 1}$ vers k' sur $B\left(0\,,\,{}^{r}/_{2}\right)$.

5) Enfin, considérons la suite $\left(k_n\right)_{n\geq 1}$ définie sur $B\left(0\,,\,{}^{r}/_{2}\right)$ par :

$$k_n = k_{\mu_n}\,.$$

on peut appliquer le théorème 2.1 et ainsi on a :

1) $\forall\, \theta \in B\left(0\,,\,\frac{r}{2}\right)$, $f(\theta) = \lim\limits_{n\to\infty} k_n(\theta)$ existe.

2) $\left(k_n\right)_{n\geq 1}$ converge vers f, uniformément sur tout borné dans $B\left(0\,,\,\frac{r}{2}\right)$.

3) f est différentiable avec $f' = k'$.

Nous noterons k la fonction limite f.

B) Nous allons démontrer que $(\mu_n)_{n\geq 1}$ converge étroitement vers une mesure de probabilité $\mu(m_0)$ telle que $k = k_{\mu(m_0)}$ sur $B(0\,,\,{}^r/_2)$.

1) Pour cela considérons pour tout θ dans $B(0\,,\,{}^r/_2)$, la loi $P(\theta,\mu_n)$ dans F_n. Alors :

$$\forall\, C>0\,,\quad \int_{\|x\|\geq C}\exp(\langle\theta,x\rangle)\,\mu_n(\mathrm{d}x) = \exp\big(k_n(\theta)\big)\cdot P(\theta,\mu_n)(\|x\|\geq C)$$

et

$$P(\theta,\mu_n)(\|x\|\geq C)\leq \frac{1}{C^2}\left(V_n(m_n)+\|m_n\|^2\right)\,,\quad \text{où } m_n = k_n'(\theta)\in J_1\,.$$

Puisque les suites $(V_n)_{n\geq 1}$ et $(k_n)_{n\geq 1}$ sont uniformément bornées, respectivement sur J_1 et $B(0\,,\,{}^r/_2)$, et puisque la suite $(m_n)_{n\geq 1}$ est bornée dans J_1, on peut conclure que pour tout $\theta\in B(0\,,\,{}^r/_2)$, la suite $\big(P(\theta,\mu_n)\big)_{n\geq 1}$ est tendue.

En appliquant le théorème de Helly, on en déduit l'existence d'une sous-suite $(\nu_k)_{k\geq 1}$ de $(\mu_n)_{n\geq 1}$ convergeant vers une mesure de probabilité dépendant de m_0 et notée $\mu(m_0)$. On a également prouvé le point suivant :

$$\forall\,\epsilon>0\,,\ \exists\, C>0:\forall\, n\geq 1 \text{ et } \forall\,\theta\in B\Big(0\,,\,\frac{r}{2}\Big)$$

on a :

$$\int_{\|X\|\geq C}\exp\left(\langle\theta\,,\,X\rangle\right)\mu_n(\mathrm{d}x)\leq\epsilon\,.$$

2) Pour démontrer la convergence étroite de $(\mu_n)_{n\geq 1}$ vers $\mu(m_0)$ nous avons besoin du lemme suivant :

LEMME. — *Soit $(\mu_n)_{n\geq 1}^{\infty}$ et μ des mesures de probabilité sur E (espace vectoriel de dimension finie d) telles que la suite $(\mu_n)_{n\geq 1}$ converge étroitement vers μ.*

Soit g une fonction continue définie sur E et à valeurs dans $\mathbb{R}$ telle que

$$\forall\,\epsilon>0\,,\ \exists\, C>0:\forall\, n\geq 1$$

on ait :

$$\int_{\|X\|\geq C}|g(x)|\mu_n(\mathrm{d}x)\leq\epsilon\,,$$

Alors :

$$\lim_{n\mapsto\infty}\int g(x)\mu_n(\mathrm{d}x) = \int g(x)\mu(\mathrm{d}x)\,.$$

Démonstration. — Tout d'abord pour toute $g : E \mapsto \mathbb{R}$, continue, il existe un ensemble au plus dénombrable de réels C tels que :

$$\mu\,\{x : g(x) = C\} > 0\,.$$

Soient alors deux réels C, C' tels que :

$$C' \geq C\,, \quad \mu\,\{x : \|x\| = C\} = \mu\,\{x : \|x\| = C'\} = 0\,.$$

Alors :

$$\lim_{n \mapsto \infty} \int_{C \leq \|x\| \leq C'} |g(x)|\mu_n(\mathrm{d}x) = \int_{C \leq \|x\| \leq C'} |g(x)|\mu(\mathrm{d}x)$$

par convergence étroite.

De plus, puisque :

$$\int_{C \leq \|x\| \leq C'} |g(x)|\mu_n(\mathrm{d}x) \leq \epsilon \quad \text{et} \quad \int_{C \leq \|x\| \leq C'} |g(x)|\mu(\mathrm{d}x) \leq \epsilon\,,$$

il existe $N(\epsilon)$ tel que $\forall\ n \geq N(\epsilon)$:

$$\begin{aligned}
&\left|\int_E g(x)\mu_n(\mathrm{d}x) - \int_E g(x)\mu(\mathrm{d}x)\right| \leq \\
&\leq \int_{\|x\| \geq C} |g(x)|\big(\mu_n(\mathrm{d}x) + \mu(\mathrm{d}x)\big) + \left|\int_{\|x\| < C} |g(x)|\big(\mu_n(\mathrm{d}x) - \mu(\mathrm{d}x)\big)\right| \\
&\leq 3\epsilon\,. \;\square
\end{aligned}$$

Appliquons le lemme à la suite $(\nu_k)_{k \geq 1}$ convergeant vers $\mu(m_0)$ et à $g(x) = \exp(\langle\theta, x\rangle)$, $\theta \in B(0\,, {}^r/_2)$.

Ainsi on obtient :

$$\lim_{n \mapsto \infty} k_n(\theta) = \log \int \exp(\langle\theta, x\rangle)\mu(m_0)(\mathrm{d}x) = k_{\mu(m_0)}(\theta)\,.$$

Puisque, sur $B(0\,, {}^r/_2)$ on a l'égalité $k_{\mu(m_0)} = k$, chaque sous-suite convergente de $(\nu_n)_{n \geq 1}$ doit converger vers la même mesure de probabilité. Donc $(\nu_n)_{n \geq 1}$ converge étroitement vers $\mu(m_0)$, mesure de probabilité sur E telle que $\Theta(\mu(m_0))$ contient $B(0\,, {}^r/_2)$ et telle que $k_{\mu(m_0)} = \lim_{n \mapsto \infty} k_{\mu_n}$ sur $B(0\,, {}^r/_2)$.

De plus, par construction, $k'_{\mu(m_0)}(0) = m_0$ est la moyenne de $\mu(m_0)$ et $k''_{\mu(m_0)}(0) = V(m_0)$ est son opérateur variance.

Puisque $V(m_0)$ est une forme quadratique strictement positive, $\mu(m_0)$ appartient à $M(E)$.

Ainsi, $\forall\, m_0 \in J$, la suite $\big(P(m_0, F_n)\big)_{n\geq 1}$ converge vers une mesure de probabilité $\mu(m_0) \in M(E)$, de moyenne m_0 et de variance $V(m_0)$.

3) Pour achever la démonstration, il faut montrer que :

$$\forall\, m_0, m_1 \text{ dans } J\,, \quad F\big(\mu(m_0)\big) = F\big(\mu(m_1)\big)\,.$$

Soit donc $m_0 \in J$, $\mu_0(m_0) = P(m_0, F_n)$ et $\mu(m_0)$ la loi limite.

Pour simplifier on note :

$$I_n(m_0) = \psi_{\mu_n(m_0)}\big(J_1(m_0)\big)\,, \quad I(m_0) = \psi_{\mu(m_0)}\big(J_1(m_0)\big)\,,$$
$$\Theta_n(m_0) = \Theta\big(\mu_n(m_0)\big)\,,$$

ainsi que :

$$\Theta(m_0) = \Theta\big(\mu(m_0)\big)\,, \quad k_n = k_{\mu_n(m_0)} \quad \text{et} \quad k = k_{\mu(m_0)}\,.$$

On sait qu'il existe $r(m_0) > 0$ tel que :

i) $B\big(0\,,\, r(m_0)\big) \subset I(m_0) \cap \left(\bigcap\limits_{n\geq 1} I_n(m_0)\right) \subset \Theta(m_0) \cap \left(\bigcap\limits_{n\geq 1} \Theta_n(m_0)\right)$.

ii) $\big(k_n\big)_{n\geq 1}$ et $\big(k'_n\big)_{n\geq 1}$ convergent respectivement vers k et k', uniformément sur toute partie bornée de $B(0\,,\, {}^{r}/_{2})$.

Soit $\eta(m_0) = \dfrac{1}{4}\, r(m_0)$.

Par un argument similaire à celui utilisé dans A) 4), on peut montrer qu'il existe $\epsilon(m_0) > 0$ tel que :

$$B(m_0) = B\big(m_0, \epsilon(m_0)\big) \subset k'\,\big(B\big(0, \eta(m_0)\big)\big) \cap \left(\bigcap_{n\geq 1} k'_n\Big(B\big(0, \eta(m_0)\big)\Big)\right)$$
$$\subset J_1(m_0)\,.$$

Pour tout m dans $B(m_0)$, on notera $\theta = \psi_{\mu(m_0)}(m)$ et $\theta_n = \psi_{\mu_n(m_0)}(m)$.

Ainsi, $\forall\, \alpha \in B\big(0, \eta(m_0)\big)$, $\theta + \alpha$ et $\theta_n + \alpha$ $(n \geq 1)$ appartiennent à $B\big(0, 2\eta(m_0)\big)$.

Donc : $\qquad k_{\mu_n(m)}(\alpha) = k_{\mu_n(m_0)}(\theta + \alpha) - k_{\mu_n(m_0)}(\theta)$

et $\qquad k_{P(\theta\,,\,\mu(m_0))}(\alpha) = k_{\mu(m_0)}(\theta + \alpha) - k_{\mu(m_0)}(\theta)$

sont bien définies sur $B\big(0,\eta(m_0)\big)$ et il est facile de voir que :

$$k_{P(\theta\,,\mu(m_0))}(\alpha) = \lim_{n\to\infty} k_{\mu_n(m)}(\alpha)\,, \quad \forall\,\alpha \in B\big(0,\eta(m_0)\big)\,.$$

Nous avons ainsi démontré que sur $B\big(0,\eta(m)\big) \cap B\big(0,\eta(m_0)\big)$ on a :

$$k_{P(\theta\,,\mu(m_0))}(\alpha) = k_{\mu(m)} \Rightarrow \mu(m) = P\left(\psi_{\mu(m_0)}(m)\,,\,\mu(m_0)\right) = P(m,F)$$

si l'on note $F = F\big(\mu(m_0)\big)$.

Donc $\forall\, m_0 \in J$, il existe $B\big(m_0\,,\,\epsilon(m_0)\big)$ telle que $\forall\, m \in B\big(m_0\,,\,\epsilon(m_0)\big)$:

$$F\big(\mu(m)\big) = F\big(\mu(m_0)\big) \quad (\Leftrightarrow \mu(m) \sim \mu(m_0))\,.$$

Soient enfin m_0 et m_1 dans J :

Si $\qquad B(m_0) \cap B(m_1) \neq \emptyset\,, \quad$ alors $\quad \mu(m_0) \sim \mu(m_1)\,.$

Donc $\mu(m) \sim \mu(m_0)$ pour tout $m \in \overline{B(m_0)}$.

Si $\qquad \overline{B(m_0)} \cap \overline{B(m_1)} = \emptyset \quad (\Leftrightarrow \|m_0 - m_1\| > \epsilon(m_0) + \epsilon(m_1))\,,$

alors on peut construire une suite $\big(x_n\big)_{n\geq 0}$ dans $[\,m_0\,,\,m_1\,] \subset J$ telle que :

$$\mu(x_n) \sim \mu(m_0) \quad \text{pour tout } n \geq 0$$

et

$$\mu(x_n) \sim \mu(m_1) \quad \text{pour } n \text{ assez grand.}$$

On définit ainsi cette suite :

– on pose $x_0 = \lambda_0 m_0 + (1-\lambda_0)m_1$, $0 < \lambda_0 < 1$ avec $\|x_0 - m_0\| = \epsilon(m_0)$, alors $\mu(x_0) \sim m_0$.

– successivement on définit :

$$x_n = \lambda_n x_{n-1} + (1-\lambda_n)m_1 \quad \text{avec} \quad \|x_n - x_{n-1}\| = \epsilon(x_{n-1})\,.$$

Alors : $\quad \mu(x_n) \sim \mu(m_0) \quad$ et $\quad x_n = (\lambda_0 \ldots \lambda_n)m_0 + (1 - \lambda_0 \ldots \lambda_n)m_1$

avec $\alpha_n = \lambda_0 \ldots \lambda_n$ tel que $0 < \alpha_{n+1} < \alpha_n$, $\forall\, n \geq 0$.

Donc pour n assez grand, $x_n \in B(m_1)$ et $\mu(x_n) \sim \mu(m_1)$.

Alors : $\qquad \mu(m_0) \sim \mu(m_1)\,, \quad \forall\, m_0, m_1 \in J\,.$

La démonstration du théorème est ainsi achevée.

Bibliographie

[1] CARTAN (H.) .— *Calcul Différentiel,*
Collection Méthodes, Hermann, Paris. (1977)

[2] LETAC (G.) .— *La réciprocité des familles exponentielles naturelles sur* $\mathbb{R}$,
C. R. Acad. Sc., Paris, **303** Série I.2 (1986) pp. 61-64

[3] LETAC (G.) et MORA (M.) .— *Natural real exponential families with cubic variance,*
(À paraître aux Ann. Statist. (mars 1990)

[4] MORA (M.) .— *Familles exponentielles et fonctions variance,*
Thèse de 3$^{\text{ième}}$ cycle. Université Paul-Sabatier, Toulouse (1986)

[5] LETAC (G.) .— *Le problème de la classification des familles exponentielles de* $\mathbb{R}^d$ *ayant une fonction variance quadratique,*
Probability on Groups IX, Lecture notes **1379** (1989) pp. 192-216. Springer

[6] MORRIS (C.N.) .— *Natural exponential families with quadratic variance functions,*
Ann. Statist., **10** (1982) pp. 65-80