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

SERGE DUBUC

Méthodes stochastiques pour la détermination de polynômes de Bernstein

Annales de l'I. H. P., section B, tome 11, nº 3 (1975), p. 203-223
<http://www.numdam.org/item?id=AIHPB_1975__11_3_203_0>

Ann. Inst. Henri Poincaré,
Vol. XI, n° 3, 1975, p. 203-223.

Section B :
Calcul des Probabilités et Statistique.

Méthodes stochastiques pour la détermination de polynômes de Bernstein

par

Serge DUBUC

RÉSUMÉ. — Nous nous intéressons à certains problèmes d'approximation et de minimisation reliés aux polynômes exponentiels de Müntz $\left\{ \sum_{\lambda \in L} c(\lambda) e^{-\lambda t} \right\}$. La considération de processus de naissance permet d'obtenir des polynômes généralisés de Bernstein utiles pour des problèmes d'approximation. Le principal problème de minimisation rencontré est de déterminer pour un ensemble fini donné, L, d'exposants, le polynôme exponentiel de Müntz $p(t) = \sum_{\lambda \in L} c(\lambda) e^{-\lambda t}$ tel que $p(t) \geq 0$ si $t \geq 0$, $\int_0^\infty p(t) dt = 1$ et $\int_0^\infty t^2 p(t) dt$ est le plus petit possible. Ici les polynômes de Laguerre interviennent de façon toute naturelle grâce à un heureux théorème de comparaison.

SUMMARY. — We consider some problems of approximation and of minimization involving exponential polynomials $\left\{ \sum_{\lambda \in L} c(\lambda) e^{-\lambda t} \right\}$. Generalized Bernstein polynomials, which are useful in approximation theory, are related to birth processes. The main problem of minimization is to find, for a given finite set, L, the Müntz polynomial $p(t) = \sum_{\lambda \in L} c(\lambda) e^{-\lambda t}$

such that $p(t) \geq 0$ if $t \geq 0$, $\int_0^\infty p(t)dt = 1$ and $\int_0^\infty t^2 p(t)dt$ is as small as possible. There is a close connection between this polynomial and Laguerre's polynomials, after a comparison theorem is proved.

Bernstein, en 1912, a trouvé une ingénieuse façon d'approcher une fonction continue sur l'intervalle-unité par une suite de polynômes [1]. Une des plus intéressantes propriétés de ces approximations est que chacun des polynômes de la suite respecte les bornes de la fonction approchée. Hirschman et Widder en 1949 [3] et Gelfond en 1950 [2] ont effectué la même démarche non plus avec les polynômes usuels mais avec les polynômes de Müntz, c'est-à-dire avec des combinaisons linéaires de puissances données de $x : \{ x^\lambda : \lambda \in \mathrm{L} \}$ alors que $\mathrm{L} = \{ \lambda_0, \lambda_1, \lambda_2, \ldots \}$, $0 = \lambda_0 < \lambda_1 \leq \lambda_2 \ldots$ et que la condition de Müntz [6] $\sum_{n=1}^{\infty} \lambda_n^{-1} = \infty$ est satisfaite ainsi que la condition $\lim_{n \to \infty} \lambda_n = \infty$. Ces auteurs ont introduit pour ce cas les polynômes généralisés de Bernstein.

On connaît une interprétation probabiliste simple des polynômes de Bernstein. C'est d'ailleurs le calcul des probabilités qui a amené Bernstein à imaginer ses polynômes. Notre première tâche sera de trouver une interprétation probabiliste des polynômes généralisés de Bernstein. Celle-ci viendra de la considération d'un processus de naissance. A cette occasion, nous serons en présence d'une classe de processus stochastiques qui permettront l'arrivée de nouveaux polynômes de Bernstein qui seront utiles surtout lorsque l'ensemble L des exposants est borné. Cette analyse réclamera la solution d'un problème de minimisation : pour un ensemble fini donné, L, d'exposants, quel est le polynôme exponentiel de Müntz $p(t) = \sum_{\lambda \in \mathrm{L}} c(\lambda) e^{-\lambda t}$ tel que $p(t) \geq 0$ si $t \geq 0$, $\int_0^\infty p(t)dt = 1$ et $\int_0^\infty t^2 p(t)dt$ est le plus petit possible. On verra qu'un problème de maximum de cet ordre fera intervenir les polynômes de Laguerre.

1. UNE CLASSE DE PROCESSUS STOCHASTIQUES

A l'aide de n variables aléatoires positives indépendantes $\xi_1, \xi_2, \ldots, \xi_n$, nous allons définir un processus stochastique $\mathrm{X}_n(t)$ dont l'espace des états sera une suite croissante donnée de nombres distincts non-négatifs

$\{ t_0, t_1, \ldots, t_n \}$. Entre l'instant 0 et ξ_1, le processus est à l'état t_0 ; entre l'instant ξ_1 et $\xi_1 + \xi_2$, le processus est à l'état $t_1, \ldots,$ à partir de l'instant $\xi_1 + \xi_2 + \ldots + \xi_n$, le processus demeure à l'état t_n. Si $p_{k,n}(t)$ représente la probabilité que le processus $X_n(t)$ soit à l'état t_k, on a que

$$p_{k,n}(t) = \Pr [0 \leq t - (\xi_1 + \xi_2 + \ldots + \xi_k) < \xi_{k+1}]$$

lorsque $k = 1, 2, \ldots, n - 1$;

$$p_{0,n}(t) = \Pr [t < \xi_1) \quad \text{et} \quad p_{n,n}(t) = \Pr [\xi_1 + \xi_2 + \ldots + \xi_n \leq t].$$

Nous prêterons à $X_n(t)$ le nom de processus de naissance généralisé. Dans ce processus, on parcourt successivement les états $t_0, t_1, \ldots, t_n$ de telle sorte que l'on demeure à l'état t_{k-1} pendant un temps aléatoire ξ_k $(k = 1, 2, \ldots, n)$.

Si les variables aléatoires $\xi_1, \xi_2, \ldots, \xi_n$ sont absolument continues par rapport à la mesure de Lebesgue et que $\Pr [\xi_k \leq t] = \int_0^t f_k(u)dx$, on peut calculer explicitement les fonctions $p_{k,n}(t)$ par intégration de divers produits de convolution :

$$\left\{ \begin{aligned} & \frac{d}{dt} p_{0,n}(t) = -f_1(t) \\ & \frac{d}{dt} p_{k,n}(t) = \left(\mathop{*}_{i=1}^{k} f_i \right)(t) - \left(\mathop{*}_{i=1}^{k+1} f_i \right)(t) \\ & \frac{d}{dt} p_{n,n}(t) = \left(\mathop{*}_{i=1}^{n} f_i \right)(t) \\ & \lim_{t \to \infty} p_{k,n}(t) = 0 \qquad (k = 0, 1, \ldots, n - 1) \\ & \lim_{t \to \infty} p_{n,n}(t) = 1 \end{aligned} \right. \tag{1}$$

Une fois fixé le processus $X_n(t)$, à une fonction réelle f définie sur $[0, \infty)$, on peut lui associer une fonction du type de Bernstein, $B_n(f)$:

$$B_n(f ; t) = \sum_{k=0}^{n} f(t_k) p_{k,n}(t) = \mathscr{E}(f(X_n(t))).$$

$B_n(f ; t)$ est l'espérance mathématique de $f(X_n(t))$.

2. CONVERGENCE DES FONCTIONS DE BERNSTEIN

Considérons une suite de processus de naissance généralisés $\{ X_n(t) \}_{n=1}^{\infty}$ où $X_n(t)$ est construit à l'aide de nombres $\{ t_{k,n} \}_{k=0}^{n}$ et de variables aléatoires $\{ \xi_{k,n} \}_{k=1}^{n}$. Soit $f(t)$ une fonction continue sur $[0, \infty)$ telle que $\lim_{t \to \infty} f(t)$ existe, nous recherchons des conditions sur les processus pour que $B_n(f ; t) = \mathscr{E}(f(X_n(t)))$ converge uniformément sur $[0, \infty)$ vers $f(t)$.

Nous introduisons d'abord la condition de concentration du tableau triangulaire des variables aléatoires $\{ \xi_{k,n} : 1 \leq k \leq n \}$: on dira que ce tableau se concentre totalement sur $[0, \infty)$ si pour tout nombre $a \geq 0$, pour tout nombre $\varepsilon > 0$ et pour tout nombre entier suffisamment grand $(n \geq N(a, \varepsilon))$, il existe un nombre entier m tel que $m \leq n$,

$$\mathscr{E}(\xi_{k,n}) < \varepsilon \qquad (1 \leq k \leq m)$$

$$\sum_{k=1}^{m} \mathscr{E}(\xi_{k,n}) \geq a$$

$$\sum_{k=1}^{m} \operatorname{Var}(\xi_{k,n}) < \varepsilon ,$$

Var ξ étant la variance de la variable ξ.

Nous dirons que le tableau triangulaire de nombres $\{ t_{k,n} : 0 \leq k \leq n \}$ est asymptotiquement bien adapté au tableau des variables aléatoires $\{ \xi_{k,n} : 1 \leq k \leq n \}$ si pour tout $\varepsilon > 0$ et si pour tout nombre n suffisamment grand $(n \geq N(\varepsilon))$,

$$|(\exp - t_{k,n}) - (\exp - \mu_{k,n})| < \varepsilon \qquad k = 0, 1, 2, \ldots, n$$

alors que $\mu_{0,n} = 0$ et $\mu_{k,n} = \sum_{i=1}^{k} \mathscr{E}(\xi_{i,n})$.

THÉORÈME 1. — Soit $X_n(t)$ une suite de processus de naissance généralisés provenant d'un tableau triangulaire de variables aléatoires obéissant à la condition de concentration sur $[0, \infty)$ et d'un tableau triangulaire de nombres asymptotiquement adapté au tableau des variables aléatoires, alors pour tout t, $X_n(t)$ converge en probabilité vers t. Plus précisément, pour tout $\varepsilon > 0$,

$$\lim_{n \to \infty} \Pr [\forall t \mid e^{-X_n(t)} - e^{-t} \mid < \varepsilon] = 1 .$$

Remarque. — La dernière assertion explique que $X_n(t)$ converge uniformément vers t si l'on utilise sur $[0, \infty)$ la structure uniforme de l'espace compact $[0, \infty]$ restreinte à l'intervalle $[0, \infty)$.

Démonstration. — Si $\{ \xi_{k,n} \}$ et $\{ t_{k,n} \}$ sont les deux tableaux responsables de la construction de la suite de processus $X_n(t)$, considérons la suite de processus $Y_n(t)$ provenant du même tableau de variables aléatoires et utilisant plutôt le tableau de nombres $\{ \mu_{k,n} : 0 \leq k \leq n \}$ où $\mu_{k,n} = \sum_{i=1}^{k} \mathscr{E}(\xi_{i,n})$.

Étudions d'abord la convergence de $Y_n(t)$. Si $t > c \geq 0$, montrons que $\lim_{n\to\infty} \Pr[Y_n(t) \leq c] = 0$. Fixons un nombre c_1, $c < c_1 < t$ et un nombre $\varepsilon > 0$. Vu l'hypothèse de concentration, pour tout n suffisamment grand, il existe un entier m tel que

$$\mu_{m,n} = \sum_{i=1}^{m} \mathscr{E}(\xi_{i,n}) \geq c_1 ,$$

$$S_{m,n} = \sum_{i=1}^{m} \operatorname{Var}(\xi_{i,n}) \leq \varepsilon$$

et

$$\mathscr{E}(\xi_{k,n}) \leq (c_1 - c) \qquad (1 \leq k \leq m).$$

Si $\mu_{k-1,n} \leq c < \mu_{k,n}$, on a que $k \leq m$.

$$\Pr[Y_n(t) \leq c] = \Pr[\xi_{1,n} + \xi_{2,n} + \dots \xi_{k,n} > t].$$

Utilisons l'inégalité de Tchebycheff :

$$\Pr[Y_n(t) \leq c] \leq \frac{\sum_{i=1}^{k} \operatorname{Var}(\xi_{i,n})}{\left(t - \sum_{i=1}^{k} \mathscr{E}(\xi_{i,n})\right)^2}$$

$$\Pr[Y_n(t) \leq c] \leq S_{k,n}(t - \mu_{k,n})^{-2}$$

or

$$\mu_{k,n} = \mu_{k-1,n} + \mathscr{E}(\xi_{k,n}) \leq c + (c_1 - c) = c_1$$

$$\Pr[Y_n(t) \leq c] \leq S_{m,n}(t - c_1)^{-2}.$$

Ce qui montre bien que $\lim_{n\to\infty} \Pr[Y_n(t) \leq c] = 0$.

Montrons maintenant que $\lim_{n\to\infty} \Pr[Y_n(t) > b] = 0$ si $0 \leq t < b$. Choisissons un nombre $\varepsilon > 0$. On peut trouver pour tout n suffisamment grand un entier m où $m \leq n$, $\mu_{m-1,n} > b$ et $S_{m,n} \leq \varepsilon$. Si $\mu_{k-2,n} \leq b < \mu_{k-1,n}$ alors $k \leq m$ et

$$\begin{aligned} \Pr[Y_n(t) > b] &= \Pr[\xi_{1,n} + \xi_{2,n} + \dots \xi_{k,n} \leq t] \\ &\leq S_{k,n}(t - \mu_{k,n})^{-2} \\ &\leq S_{m,n}(t - b)^{-2}. \end{aligned}$$

Ce qui montre que $\lim_{n\to\infty} \Pr [Y_n(t) > b] = 0$. Les estimés obtenus sur les probabilités joints au fait que $t \to Y_n(t)$ ne décroît pas montrent que

$$\lim_{n\to\infty} \Pr [\forall t \mid e^{-Y_n(t)} - e^{-t} \mid < \varepsilon] = 1 .$$

Si l'on revient à la suite $X_n(t)$, la condition que $\{ t_{k,n} : 0 \leq k \leq n, n \geq 1 \}$ est asymptotiquement adapté au tableau $\{ \xi_{k,n} \}$ permet d'affirmer que

$$\lim_{n\to\infty} \Pr [\forall t \mid e^{-Y_n(t)} - e^{-X_n(t)} \mid < \varepsilon] = 1 .$$

D'où la conclusion du théorème 1.

COROLLAIRE. — Si $X_n(t)$ est une suite de processus de naissance généralisés obéissant aux hypothèses du théorème précédent et si $f(t)$ est une fonction continue sur $[0, \infty)$ telle que $\lim_{t\to\infty} f(t)$ existe, alors $\mathscr{E}(f(X_n(t)))$ converge vers $f(t)$ uniformément en t.

3. POLYNÔMES GÉNÉRALISÉS DE BERNSTEIN

Dans cette section, nous rappelons les polynômes généralisés de Bernstein tels que définis par Hirschman, Widder et Gelfond. Nous indiquerons comment le corollaire du théorème 1 permet d'englober une partie de leurs résultats. Présentons d'abord des préliminaires. Si $L = \{ \lambda_1, \lambda_2, \ldots, \lambda_m \}$ est un ensemble de nombres réels permettant la répétition, une somme d'exponentielles (ou un polynôme exponentiel de Müntz) à paramètres dans L est la totalité P(L) des combinaisons linéaires des fonctions $t \to t^{r-1}e^{-\lambda t}$ où $\lambda \in L$ et où r est un entier positif qui ne dépasse pas le nombre de fois où λ est répété dans L.

$$P^+(L) = \{ p \in P(L) : t \geq 0 \Rightarrow p(t) \geq 0 \} .$$

On notera par L^* l'augmentation de 0 à L quand $0 \notin L$: $L^* = L$ si $0 \in L$ et $L^* = L \cup \{ 0 \}$ si $0 \notin L$. Si $L = \{ \lambda_1, \ldots, \lambda_m \}$ et si $M = \{ \mu_1, \ldots, \mu_n \}$, on entendra par $L(+)M = \{ \lambda_1, \lambda_2, \ldots, \lambda_m, \mu_1, \mu_2, \ldots, \mu_n \}$. De façon analogue, on définit $L_1(+)L_2(+) \ldots (+)L_k$.

Considérons comme tantôt un tableau de variables aléatoires indépendantes $\{ \xi_{k,n} : 1 \leq k \leq n \}$ et un tableau de nombres $\{ t_{k,n} \}_{k=0}^n$. Nous dégagerons quelques propriétés des fonctions $p_{k,n}(t) = \Pr [X_n(t) = t_{k,n}]$ si

$$\Pr [\xi_{k,n} > t] = \int_t^\infty f_{k,n}(u)du$$

alors que $f_{k,n}$ est une somme d'exponentielles.

LEMME 2. — Si $f \in P(L)$ et $g \in P(M)$, alors le produit de convolution de f avec g appartient à $P(L(+)M)$.

Démonstration. — Le cas important à traiter est lorsque

$$\lambda_1 = \lambda_2 = \ldots = \lambda_m = \lambda, \qquad \mu_1 = \mu_2 = \ldots = \mu_n = \mu$$

et $f(t) = t^{m-1}e^{-\lambda t}$, $g(t) = t^{n-1}e^{-\mu t}$. Si $\lambda = \mu$,

$$(f*g)(t) = \left(\int_0^t (t-u)^{m-1}u^{n-1}du\right)e^{-\lambda t} = B(m, n)t^{m+n-1}e^{-\lambda t}.$$

Si $\lambda \neq \mu$,

$$\begin{aligned}(f*g)(t) &= \int_0^t (t-u)^{m-1}e^{-\lambda(t-u)}u^{n-1}e^{-\mu u}du \\ &= \frac{\partial^{m+n-2}}{\partial\lambda^{m-1}\partial\mu^{n-1}}\int_0^t e^{-\lambda(t-u)}e^{-\mu u}du \\ &= \frac{\partial^{m+n-2}}{\partial\lambda^{m-1}\partial\mu^{n-1}}\left(\frac{e^{-\mu x}}{\lambda-\mu} + \frac{e^{-\lambda x}}{\mu-\lambda}\right).\end{aligned}$$

On voit bien que cette dernière expression appartient à $P(L(+)M)$.

LEMME 3. — Si les fonctions de densité $f_{k,n}$ des variables $\xi_{k,n}$ appartiennent à $P(L'_{k,n})$, alors $p_{k,n}(t)$ appartient à

$$P\left(\left(\sum_{i=1}^{k+1}\right)L'_{i,n}\right) \qquad k = 0, 1, \ldots, n-1$$

et $p_{n,n}(t)$ appartient à

$$P\left(\left(\left(\sum_{i=1}^{n}\right)L'_{i,n}\right)^*\right).$$

Démonstration. — Les formules (1) de la section 1 et le lemme 2 donnent que $\frac{d}{dt}p_{k,n}(t)$ appartient à

$$P\left(\left(\sum_{i=1}^{k+1}\right)L'_{i,n}\right)$$

lorsque $k \leq n-1$. Dans ce cas, $\lim_{t\to\infty} p_{k,n}(t) = 0$; d'où la conclusion de ce lemme. D'autre part

$$\frac{d}{dt}p_{n,n}(t) \in P\left(\left(\sum_{i=1}^{n}\right)L'_{i,n}\right)$$

et l'intégration de cette fonction de t à l'infini fait apparaître une constante.

Si $\{ \lambda_k \}_{k=0}^{\infty}$ est une suite où $0 = \lambda_0$, $\lambda_k > 0$ $(k \geq 1)$, utilisons comme fonction de densité de $\xi_{k,n}$ la fonction $\lambda_{n+1-k}e^{-t\lambda_{n+1-k}}$. Nous spécifierons tantôt le tableau de nombres $\{ t_{k,n} \}_{k=0}^{n}$. Le processus $X_n(t)$ est connu sous le nom de processus de naissance (voir par exemple Parzen [7]). Le lemme 3 permet de dire que $p_{k,n}(t)$ appartient à $P(L_{k,n})$ où

$$L_{k,n} = \{ \lambda_n, \lambda_{n-1}, \ldots, \lambda_{n-k} \}.$$

Hirschman et Widder [3], pour définir des polynômes généralisés de Bernstein lorsque les λ_k sont tous distincts, ont introduit des fonctions $H_{k,n}(t)$: $H_{n,n}(t) = e^{-\lambda_n t}$ et pour $k = 0, 1, \ldots,$ ou $n - 1$ $H_{k,n}(t)$ est l'unique solution $y(t)$ à l'équation différentielle

$$\left[\left(1 + \lambda_{k+1}^{-1}\frac{d}{dt}\right)\left(1 + \lambda_{k+2}^{-1}\frac{d}{dt}\right) \ldots \left(1 + \lambda_n^{-1}\frac{d}{dt}\right)\right] y(t) = e^{-\lambda_k t}$$
$$y(0) = y'(0) = \ldots\ y^{(n-k-1)}(0) = 0.$$

Ils ont posé $\sigma_{0,n} = \infty$ et

$$\sigma_{k,n} = \sum_{i=k}^{n} \lambda_i^{-1}$$

et ils ont proposé le polynôme généralisé de Bernstein

$$\sum_{k=0}^{n} f(\sigma_{k,n}) H_{k,n}(t) = H_n(f\ ;\ t).$$

Gelfond introduit des polynômes généralisés de Bernstein de façon légèrement différente et qui généralisent de façon directe les polynômes de Bernstein. Son approche, qui est exposée dans le volume de Lorentz [5], utilise les différences finies de la fonction $\lambda \to e^{-\lambda t}$. Il obtient des fonctions $p_{k,n}^*(t)$ de la façon suivante :

$$p_{k,n}^*(t) = \frac{(-1)^{n-k}}{2\pi i} \lambda_{k+1}\lambda_{k+2} \ldots \lambda_n \int_C \frac{e^{-zt}dz}{(z - \lambda_k)(z - \lambda_{k+1}) \ldots (z - \lambda_n)} \qquad k = 0, 1, \ldots, n-1$$

où C est une courbe simple fermée orientée positivement du plan complexe qui contient les nombres $\lambda_0, \lambda_1, \ldots, \lambda_n$ et $p_{n,n}^*(t) = e^{-\lambda_n t}$. Il considère les abscisses

$$A_{k,n} = \frac{1}{\lambda_1} \sum_{i=k+1}^{n} \log(1 - \lambda_1/\lambda_i) \qquad 0 \leq k < n$$
$$A_{n,n} = 0$$

et il propose le polynôme

$$B_n^*(f\ ;\ t) = \sum_{k=0}^{n} f(A_{k,n}) p_{k,n}^*(t) .$$

Le choix d'abscisses $\{ A_{k,n} \}$ peut paraître bizarre ; c'est pourtant le choix convenable pour que $B_n^*(f\ ;\ t) = f(t)$ si $f(t) = e^{-\lambda_1 t}$.

On peut voir que $H_{k,n}(t) = p_{k,n}^*(t) = p_{n-k,n}(t)$. Il suffit de prendre la transformée de Laplace de chacune de ces fonctions qui est dans tous les cas

$$\frac{1}{s + \lambda_k} \prod_{i=k+1}^{n} (\lambda_i/(s + \lambda_i)) \qquad (0 \leq k < n)$$

et $(s + \lambda_n)^{-1}$ si $k = n$. Pour le calcul de la transformée de Laplace de $p_{n-k,n}(t)$, on se sert des formules (1).

Si $X_n^*(t)$, $X_n^{**}(t)$ représentent les processus de naissance généralisés utilisant les tableaux triangulaires de nombres $\{ A_{n-k,n} \}_{k=0}^n$ et $\{ \sigma_{n-k,n} \}_{k=0}^n$, on obtient que $\mathscr{E}(f(X_n^*(t))) = B_n^*(f\ ;\ t)$ et $\mathscr{E}(f(X_n^{**}(t))) = H_n(f\ ;\ t)$.

Il est facile de vérifier que les hypothèses du théorème 1 sont satisfaites si $\{ \lambda_n \}_{n=1}^{\infty}$ est une suite qui croît sans borne alors que $\sum_{\lambda=1}^{\infty} \frac{1}{\lambda_n} = \infty$, ceci pour les processus $X_n^*(t)$ et $X_n^{**}(t)$. Si $\varepsilon > 0$ et $a > 0$ sont donnés, on peut trouver un entier N tel que $\lambda_k > \frac{1}{\varepsilon}$ si $k \geq N$. Puisque $\sum_{n=1}^{\infty} \frac{1}{\lambda_n} = \infty$, on peut trouver un entier N, tel que $\sum_{n=N}^{N_1} \frac{1}{\lambda_n} \geq A$. Si $n \geq N_1$, on peut trouver un entier m tel que $m \leq n$ et

$$a \leq \sum_{k=1}^{m} \mathscr{E}(\xi_{k,n}) < a + \varepsilon$$

c'est-à-dire que

$$a \leq \sum_{k=0}^{m-1} \frac{1}{\lambda_{n-k}} < a + \varepsilon .$$

Dans ce cas,

$$\sum_{k=1}^{m} \text{Var } \xi_{k,n} = \sum_{k=0}^{m-1} \frac{1}{\lambda_{n-k}^2}$$

$$\sum_{k=1}^{m} \text{Var } \xi_{k,n} \leq \frac{1}{\lambda_{n-m+1}} \sum_{k=0}^{m-1} \frac{1}{\lambda_{n-k}}$$

$$\sum_{k=1}^{m} \text{Var } \xi_{k,n} \leq \varepsilon(a + \varepsilon) .$$

Les variables $\{ \xi_{k,n} \}$ sont donc totalement concentrées sur $[0, \infty)$. De façon semblable, on montre que les tableaux $\{ \sigma_{n-k,n} \}$ et $\{ a_{n-k,n} \}$ sont adaptés aux variables $\{ \xi_{k,n} \}$.

4. UN PROBLÈME DE MINIMISATION POUR LES SYSTÈMES DE TCHEBYCHEFF

Avant d'appliquer une autre fois le théorème 1, nous allons traiter d'un problème auxiliaire, celui de caractériser pour un ensemble donné d'exposants positifs $L = \{ \lambda_0, \lambda_1, \ldots, \lambda_n \}$ la fonction de densité $p(t)$ de $P^+(L)$ $\left(\int_0^\infty p(t)dt = 1 \right)$ telle que $\int_0^\infty t^k p(t)dt$ est minimale (où $k > 0$). Ce problème est un cas particulier d'un problème plus général que nous allons analyser pour les systèmes de Tchebycheff de fonctions. Si I est un intervalle de $\mathbf{R}$, un système de Tchebycheff de degré n sur I est un sous-espace vectoriel P de dimension $n + 1$ de l'espace vectoriel C(I) des fonctions continues de I dans $\mathbf{R}$ tel que pour tout $p \in P$, le nombre de racines à l'équation $p(t) = 0$, $t \in I$, ne dépasse pas n. Comme l'indiquent Karlin et Studden à la page 9 de [*4*], $P(\lambda_0, \lambda_1, \ldots, \lambda_n)$ est un système de Tchebycheff sur $(-\infty, \infty)$.

Si P est un système de Tchebycheff sur $[0, \infty)$, nous dirons que P est un système de Tchebycheff restreint s'il existe une fonction u de P telle que

a) pour tout $v \in P$, $\lim_{t \to \infty} \frac{v(t)}{u(t)}$ existe,

b) $\left\{ v : \lim_{t \to \infty} \frac{v(t)}{u(t)} = 0 \right\}$ est un système de Tchebycheff de degré $n - 1$ sur $[0, \infty)$.

Nous avons repris une définition due à Karlin-Studden ([*4*], p. 148). Lorsque P est un système de Tchebycheff restreint, nous introduisons la fonctionnelle $p(\infty *) = \lim\limits_{t \to \infty} \dfrac{p(t)}{u(t)}$. A une constante multiplicative près, cette fonctionnelle est uniquement déterminée. Encore ici, $P(\lambda_0, \lambda_1, \ldots, \lambda_n)$ est un système de Tchebycheff restreint sur $[0, \infty)$; si $\lambda_0 \leq \lambda_1 \leq \ldots \leq \lambda_n$ et si le nombre λ_0 est présent r fois dans la suite, on peut prendre pour fonction $u(t)$ la fonction $t^{r-1}e^{-\lambda_0 t}$.

Nous désignerons par M(P), l'espace des moments de P, l'ensemble des fonctionnelles linéaires de P dans **R**. Si P est un système de Tchebycheff restreint sur $[0, \infty)$, si $T \in M(P)$, une représentation discrète de T est la donnée de nombres $\rho_\infty, \rho_0, \rho_1, \ldots, \rho_s, t_1, t_2, \ldots, t_s$ où $\{ t_1, t_2, \ldots, t_s \}$ sont des nombres distincts de $(0, \infty)$, $\rho_i \neq 0$ si $i = 1, 2, \ldots, s$ et $\forall p \in P$

$$t(p) = \rho_0 p(0) + \rho_\infty p(\infty *) + \sum_{i=1}^{s} \rho_i p(t_i) .$$

Si $s = 0$, on fait la convention que $\sum_{i=1}^{0} \rho_i p(t_i) = 0$. Si $A = \{ t_1, t_2, \ldots, t_s \}$, on définit le degré de la représentation et les abscisses de la représentation comme suit :

a) $\rho_0 \neq 0$, $\rho_\infty \neq 0$, le degré est $2s+2$ et les abscisses sont $\{ 0, \infty \} \cup A$
b) $\rho_0 \neq 0$, $\rho_\infty = 0$, le degré est $2s+1$ et les abscisses sont $\{ 0 \} \cup A$
c) $\rho_0 = 0$, $\rho_\infty \neq 0$, le degré est $2s+1$ et les abscisses sont $A \cup \{ \infty \}$
d) $\rho_0 = 0$, $\rho_\infty = 0$, le degré est $2s$ et les abscisses sont les nombres de A.

On dit que la représentation est positive si tous les nombres ρ_i sont non-négatifs. D'autre part, on dira que le degré de $T \in M(P)$ est k s'il existe une représentation de degré k de T mais qu'il n'existe pas de représentation de degré inférieur à k pour T.

Si P^+ est la totalité des fonctions non-négatives sur I de P, $M^+(P)$ sera l'ensemble des fonctionnelles T telles que $\forall p \in P^+$, $t(p) \geq 0$, ce sont les fonctionnelles non-négatives sur le cône convexe P^+. Les fonctionnelles qui sont à la frontière de $M^+(P)$ jouissent d'une propriété remarquable.

THÉORÈME 4. — Si P est un système de Tchebycheff restreint de degré n sur $[0, \infty)$, si T est une fonctionnelle à la frontière de $M^+(P)$, alors le degré de T ne dépasse pas n et T admet une seule représentation discrète de degré inférieur ou égal à n et celle-ci est une représentation positive. Si A est l'ensemble des abscisses de la représentation, $\{ p : p \geq 0, T(p) = 0 \} = \{ p : p \geq 0, p(A) = 0 \}$. En particulier, si le degré de T est n, à une constante multiplicative près, il existe un seul p tel que $p \geq 0$ et $T(p) = 0$.

Démonstration. — La première phrase de ce théorème se démontre comme le théorème 2.1 du chapitre II (p. 42) du volume de Karlin-Studden [*4*]. La deuxième phrase se démontre facilement. La troisième phrase est une conséquence de l'adaptation aux systèmes de Tchebycheff restreints du théorème 4.2 du chapitre I (p. 23) de Karlin-Studden [*4*].

Remarque. — Une fonctionnelle T à la frontière de $M^+(P)$ est une fonctionnelle de $M^+(P)$ qui annule au moins un polynôme non-négatif p, qui n'est pas identiquement nul.

Nous arrivons à la principale proposition de cette section.

THÉORÈME 5. — Soit P un système de Tchebycheff restreint de degré n sur $[0, \infty)$, soit μ une mesure positive sur $[0, \infty)$ et $\sigma(t)$ une fonction croissante sur $[0, \infty)$ telles que

a) le support de la mesure μ contient plus de $\frac{n}{2}$ points

b) pour toute fonction p de P,

$$\int_0^\infty (|\sigma(t)| + 1)\,|p(t)|\,d\mu(t) < \infty,$$

alors le problème de déterminer la valeur

$$m = \inf\left\{ \int_0^\infty p(t)\sigma(t)d\mu(t) : p \in P^+,\ \int_0^\infty p(t)d\mu(t) = 1 \right\}$$

est équivalent à la recherche d'une valeur m_1 telle que la fonctionnelle

$$\int_0^\infty p(t)(\sigma(t) - m_1)d\mu(t)$$

admet une représentation discrète positive de degré inférieur ou égal à n. On a que $m_1 = m$; le degré de la représentation est nécessairement à n. Pour tout abscisse c de la représentation $\sigma(c) > m$. Le polynôme positif où est atteint le minimum m est le polynôme p tel que $\int_0^\infty p(t)d\mu(t) = 1$ et dont les zéros sont les abscisses de la représentation discrète.

Démonstration. — Dire que $\int_0^\infty p(t)(\sigma(t) - m_1)d\mu(t)$ admet une représentation discrète positive, assure que pour tout polynôme positif p,

$$\int_0^\infty p(t)\sigma(t)d\mu(t) \geqslant m_1 \int_0^\infty p(t)d\mu(t).$$

Si A est l'énumération des abscisses de cette représentation discrète dont le degré ne dépasserait pas n, alors on peut trouver un polynôme positif p_0 qui s'annule sur A et tel que $\int_0^\infty p_0(t)d\mu(t) = 1$. D'où

$$\int_0^\infty p_0(t)(\sigma(t) - m_1)d\mu(t) = 0 \qquad \text{et} \qquad m = m_1 .$$

Posons $\alpha = \inf \{ t : \sigma(t) > m \}$. Montrons que selon la parité de n il y aura $n/2$ ou $(n + 1)/2$ abscisses de A dans l'intervalle $(\alpha, \infty]$. Si n est pair et s'il n'y avait pas $\frac{n}{2}$ abscisses de A dans l'intervalle (α, ∞), on pourrait alors trouver un polynôme p de P tel que

$$\begin{array}{lll} p(t) < 0 & \text{si} & t < \alpha \\ p(\alpha) = 0 & & \\ p(t) > 0 & \text{si} & t \notin \mathrm{A} \cap (\alpha, \infty] \\ p(c) = 0 & \text{si} & c \in \mathrm{A} \cap (\alpha, \infty] . \end{array}$$

Vu les contraintes de signes sur p, on a que $\int_0^\infty p(t)(\sigma(t) - m)d\mu(t) \geq 0$. Or la donnée de la représentation discrète assure que

$$\int_0^\infty p(t)(\sigma(t) - m)d\mu(t) \leq 0 .$$

D'où le support de la mesure μ serait concentré dans $\{ t : p(t)(\sigma(t) - m) = 0 \}$, ce qui contredit l'hypothèse que le support de μ contient plus de $\frac{n}{2}$ points. Lorsque n est impair, on obtient avec la même argumentation que les abscisses de la représentation sont constitués de ∞ et de $\frac{n-1}{2}$ nombres de l'intervalle (α, ∞). Du même coup, on a que le degré de la représentation est égal à n dans les deux cas. La dernière partie du théorème 5 provient directement du théorème 4. C. Q. F. D.

La solution au problème de minimum qui vient d'être présentée revient donc à résoudre un problème non-linéaire de $(n + 1)$ équations à $(n + 1)$ inconnues. Si $\{ p_i(t) \}_{i=0}^n$ est une base de P, et, si n est pair, les équations à résoudre sont

$$\int_0^\infty p_i(t)\sigma(t)d\mu(t) = m \int_0^\infty p_i(t)d\mu(t) + \sum_{j=1}^{n/2} \rho_j p_i(t_j) \qquad i = 0, 1, \ldots, n$$

les inconnues étant m, $\rho_j > 0$, $t_j > 0$, $j = 1, 2, \ldots, n/2$. Lorsque n est impair, les équations sont plutôt :

$$\int_0^\infty p_i(t)(\sigma(t) - m)d\mu(t) = \rho_\infty p_i(\infty *) + \sum_{j=1}^{(n-1)/2} \rho_j p_i(t_j) .$$

Ceci revient à déterminer des multiplicateurs de Kuhn-Tucker.

THÉORÈME 6 (Théorème de comparaison). — Soient $\{ f_0, f_1, \ldots, f_n \}$ $n+1$ fonctions continues sur $[0, \infty)$ telles que $\lim_{t\to\infty} f_i(t)/f_{i-1}(t) = 0$, $i = 1, 2, \ldots, n$, on désigne par P_0, P_1 et P_2 les sous-espaces vectoriels de fonctions engendrés respectivement par $\{ f_0, f_1, \ldots, f_n \}$, $\{ f_0, f_1, \ldots, f_{n-1} \}$ et $\{ f_1, f_2, \ldots, f_n \}$. On suppose que ces sous-espaces sont des systèmes de Tchebycheff restreints. Soient de plus μ une mesure positive sur $[0, \infty)$ et $\sigma(t)$ une fonction croissante remplissant les conditions du théorème 5 pour P, si

$$m_i = \inf \left\{ \int_0^\infty f(t)\sigma(t)d\mu(t) : f \in P_i^+ \text{ et } \int_0^\infty f(t)d\mu(t) = 1 \right\}$$

alors $m_2 \leq m_1$.

Démonstration. — Traitons d'abord du cas où n est impair. Soit $g(t)$ la fonction de P_0^+ telle que $\int_0^\infty g(t)d\mu(t) = 1$ et $\int_0^\infty g(t)\sigma(t)d\mu(t) = m_0$. Vu le théorème 5 et l'imparité de n, le point à l'infini est un abscisse de la représentation discrète de $f \to \int_0^\infty f(t)(\sigma(t) - m_0)d\mu(t)$ et ainsi $g(\infty *) = 0$. Mais

$$g(t) = \sum_{i=0}^{n} c_i f_i(t)$$

et

$$g(\infty *) = \lim_{t\to\infty} g(t)/f_0(t) = c_0 .$$

D'où $g \in P_2$ et $m_2 = m$. D'autre part, il est clair que $m \leq m_1$. D'où la conclusion $m_2 \leq m_1$.

Si n est pair, soit $h(t)$ la fonction de P_1^+ telle que $\int_0^\infty h(t)d\mu(t) = 1$ et $\int_0^\infty h(t)\sigma(t)d\mu(t) = m_1$. Pour le même motif que tantôt, on obtient que $h(\infty *) = 0$ et que $h(t) = \sum_{i=0}^{n-1} d_i f_i(t)$ avec $d_0 = 0$. Ainsi $h \in P_2$ et

$$m_2 \leq \int_0^\infty h(t)\sigma(t)d\mu(t) = m_1 . \qquad \text{C. Q. F. D.}$$

5. SOMMES D'EXPONENTIELLES DONT LE PREMIER MOMENT EST PETIT

Revenons aux sommes d'exponentielles. Si $L = \{ \lambda_1, \lambda_2, \ldots, \lambda_n \}$, $0 < \lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n$, on a déjà noté que P(L) est un système restreint de Tchebycheff sur $[0, \infty)$. Si $f_i(t) = t^{r_i - 1}e^{-\lambda_i t}$ où r_i est le nombre de fois où λ_i apparaît dans la suite $\{ \lambda_i, \lambda_{i+1}, \ldots, \lambda_n \}$, on a que

$$\lim_{t \to \infty} \frac{f_{i+1}(t)}{f_i(t)} = 0 .$$

On peut donc appliquer les résultats de la section précédente aux sommes d'exponentielles. Posons ici

$$m(L) = \inf \left\{ \int_0^\infty tp(t)dt : p \in P^+(L) \text{ et } \int_0^\infty p(t)dt = 1 \right\}$$

$m(L)$ est la plus petite moyenne que l'on peut obtenir en utilisant une fonction de densité de probabilité sur $(0, \infty)$ qui soit dans P(L).

Le théorème 5 fait voir que la recherche de $m(L)$ et du polynôme $p(t)$ minimisant revient à trouver le nombre m ($m(L)$) et des nombres positifs $x_1, x_2, \ldots, x_r$ si $n = 2r + 1$ ou $2r + 2$ tel que

$$\frac{r_i!}{\lambda_i^{r_i+1}} = \frac{m(r_i - 1)!}{\lambda_i^{r_i}} + \sum_{j=1}^{r} \rho_j x_j^{r_i - 1} e^{-\lambda_i x_j}$$

$$\begin{array}{ll} i = 1, 2, \ldots, n & \text{si } n \text{ est impair} \\ i = 2, 3, \ldots, n & \text{si } n \text{ est pair.} \end{array}$$

Et l'on obtient que $x_j > m(L)$. La quantité $m(L)$ est une fonction apparemment compliquée de L. Déterminons de façon plus précise $m_n = m(L)$ lorsque $L = \{ \lambda_1, \lambda_2, \ldots, \lambda_n \}$ avec $\lambda_1 = \lambda_2 = \ldots = \lambda_n = 1$.

THÉORÈME 7. — Si $p_n(t)$ est le n^e polynôme de Laguerre et si ξ_n est la plus petite racine de p_n, on a que pour tout polynôme $q(t)$ de degré inférieur ou égal à $2n - 1$ tel que $q(t) \geq 0$ lorsque $t \geq 0$, on a que

$$\int_0^\infty tq(t)e^{-t}dt \geq \xi_n \int_0^\infty q(t)e^{-1}dt ,$$

l'égalité n'étant atteinte que si $q(t)$ est un multiple scalaire de

$$\left(\frac{p_n(t)}{t - \xi_n} \right)^2 .$$

Démonstration. — Soient $\xi_{1,n}, \xi_{2,n}, \ldots, \xi_{n,n}$ les n zéros, ordonnés par ordre croissant, de $p_n(t)$; par la méthode de quadrature de Gauss (cf. Szegö [9] par exemple), on sait que l'on peut trouver des nombres positifs $H_1, H_2, \ldots, H_n$ (les nombres de Christoffell) tels que

$$\int_0^\infty q(t)e^{-t}dt = \sum_{i=1}^n H_i q(\xi_{i,n})$$

si le degré de q ne dépasse pas $2n - 1$. Si le degré de q ne dépasse pas $2n - 2$ on a donc que

$$\int_0^\infty q(t)te^{-t}dt = \sum_{i=1}^n H_i\xi_{i,n}q(\xi_{i,n})$$

et

$$\int_0^\infty q(t)(t - \xi_n)e^{-t}dt = \sum_{i=2}^n H_i(\xi_{i,n} - \xi_n)q(\xi_{i,n}) .$$

Ainsi $m_{2n-1} = \xi_n$ et le polynôme extrémal est alors

$$\prod_{i=2}^n (t - \xi_{in})^2 \qquad \text{ou} \qquad \left(\frac{p_n(t)}{t - \xi_n}\right)^2$$

à une constante multiplicative près. D'autre part $m_{2n} = m_{2n-1}$ et donne le même polynôme extrémal.

THÉORÈME 8. — Si $L = \{ \lambda_1, \lambda_2, \ldots, \lambda_{2n-1} \}$ et si ξ_n est la plus petite racine du n^e polynôme de Laguerre

$$\xi_n/\lambda_{2n-1} \leq m(L) \leq \xi_n/\lambda_1 .$$

Démonstration. — On considère la suite

$$\lambda_1, \lambda_1, \ldots, \lambda_1, \qquad \lambda_2, \ldots, \lambda_{2n-1}, \qquad \lambda_{2n-1}, \ldots, \lambda_{2n-1}$$

où λ_1 est répété initialement $(2n - 1)$ fois et λ_{2n-1} est répété finalement $(2n - 1)$ fois et l'on considère L_k formé de $2n - 1$ exposante consécutifs issus de la dernière suite ($k = 1, 2, \ldots, 4n - 3$). En appliquant le théorème 6, on a que $m(L_{k+1}) \leq m(L_k)$. D'où $m(L_{4n-3}) \leq m(L) \leq m(L_1)$. Mais

$$m(L_1) = \frac{\xi_n}{\lambda_1} \qquad \text{et} \qquad m(L_{4n-3}) = \frac{\xi_n}{\lambda_{2n-1}} .$$

COROLLAIRE. — Si $\{ \lambda_k \}_{k=1}^\infty$ est une suite croissante, si

$$L_n = \{ \lambda_1, \lambda_2, \ldots, \lambda_{2n-1} \},$$

alors il existe deux constantes universelles A et B telles que

$$\frac{A}{n\lambda_1} \leq m(L_n) \leq \frac{B}{n\lambda_{2n-1}}.$$

En effet, le plus petit zéro du n^e polynôme de Laguerre est de l'ordre de $\frac{1}{n}$ (cf. Szegö [9], p. 129).

A titre de curiosité, citons une majoration de la variance de la loi de probabilité de densité $p(t)$ telle que $\int_0^\infty tp(t)dt = m(L)$.

THÉORÈME 9. — Si $q(t)$ est le polynôme de Müntz de $P^+(\lambda_1, \lambda_2, \ldots, \lambda_{2n+1})$ qui minimise la quantité $\int_0^\infty tq(t)dt$ sous la contrainte $\int_0^\infty q(t)dt = 1$, si m est ce minimum et si $\sum_{j=1}^n \rho_j p(x_j)$ est la représentation canonique de la fonctionnelle

$$p \to \int_0^\infty (t - m)p(t)dt,$$

alors

$$\int_0^\infty t^2 q(t)dt - \mu^2 \leq \sum_{j=1}^n \rho_j.$$

Démonstration. — On remarque que

$$\int_0^\infty t^2 q(t)dt = \int_0^\infty t \int_t^\infty q(u)dudt.$$

Or $\int_t^\infty q(u)du$ appartient à $P(\lambda_1, \lambda_2, \ldots, \lambda_{2n+1})$.

$$\int_0^\infty t \int_t^\infty q(u)dudt = \mu \int_0^\infty \int_t^\infty q(u)du + \sum_{j=1}^n \rho_j \int_{x_j}^\infty q(u)du$$

$$\int_0^\infty t^2 q(t)dt = \mu^2 + \sum_{j=1}^n \rho_j \int_{x_j}^\infty q(u)du \leq \mu^2 + \sum_{j=1}^n \rho_j.$$

6. SOMMES D'EXPONENTIELLES DONT LE SECOND MOMENT EST PETIT

Si L est une collection d'exposants positifs, nous posons

$$s(L) = \inf\left\{ \int_0^\infty t^2 p(t)dt : p \in P^+(L) \text{ et } \int_0^\infty p(t)dt = 1 \right\}.$$

$s(\mathrm{L})$ est le plus petit second moment que l'on peut obtenir en utilisant une densité de probabilité sur $(0, \infty)$ qui soit dans $\mathrm{P}(\mathrm{L})$.

THÉORÈME 10. — Si $p_n(t)$ et $p_{n+1}(t)$ sont les polynômes de Laguerre de degré n et $n+1$, si θ_n est la plus petite racine positive du polynôme $p_{n+1}(t)p_n(-t) - p_{n+1}(-t)p_n(t)$, on a que pour tout polynôme $q(t)$ de degré inférieur ou égal à $(2n-1)$, non-négatif sur $(0, \infty)$,

$$\int_0^\infty t^2 q(t)e^{-t}dt \geq \theta_n^2 \int_0^\infty q(t)e^{-t}dt .$$

L'égalité n'a lieu que pour un multiple de

$$(p_{n+1}(t)p_n(\theta_n) - p_{n+1}(\theta_n)p_n(t))^2/(t^2 - \theta_n^2)^2 .$$

Démonstration. — Si on pose $\mathrm{L}_n = \{ 1, 1, \ldots, 1 \}$, 1 étant présent $2n$ fois dans L, on veut démontrer que $s(\mathrm{L}_n) = \theta_n^2$. Soit $q(t)$ le polynôme de degré inférieur ou égal à $2n-1$ tel que $q(t) \geq 0$ lorsque $t \in (0, \infty)$,

$$\int_0^\infty q(t)e^{-t}dt = 1 \qquad \text{et} \qquad \int_0^\infty t^2 q(t)e^{-t}dt = s(\mathrm{L}_n).$$

Par le théorème 5, $q(t)e^{-t}|_{t=\infty^*} = 0$, c'est-à-dire que le degré de q n'est pas égal à $(2n-1)$, et $q(t)$ est un polynôme non-négatif qui s'annule à $(n-1)$ endroits distincts. $q(t)$ est donc le carré d'un polynôme $p(t)$ de degré $(n-1)$. Ainsi

$$s(\mathrm{L}_n) = \inf \left\{ \int_0^\infty t^2 p^2(t)e^{-t}dt : \int_0^\infty p^2(t)e^{-t}dt = 1 \text{ et } \deg p \leq n-1 \right\}.$$

La méthode des multiplicateurs de Lagrange permet d'affirmer qu'il existe un scalaire λ tel que pour tout polynôme $r(t)$ de degré inférieur ou égal à $(n-1)$

$$\int_0^\infty t^2 p(t)r(t)e^{-t}dt = \lambda \int_0^\infty p(t)r(t)e^{-t}dt$$

ou

$$\int_0^\infty (t^2 - \lambda)p(t)r(t)e^{-t}dt = 0 .$$

Autrement dit, pour le produit scalaire $\langle f, g \rangle = \int_0^\infty f(t)g(t)e^{-t}dt$, la fonction $(t^2 - \lambda)p(t)$ est orthogonale à tout polynôme de degré inférieur ou égal à $n-1$. On peut donc trouver deux constantes réelles α et β telles que

$$(t^2 - \lambda)p(t) = \alpha p_{n+1}(t) + \beta p_n(t)$$

vu le développement de Fourier-Laguerre du membre à gauche. On peut remarquer que $\lambda = s(L_n)$ et l'on aura que

$$\alpha p_{n+1}(\sqrt{\lambda}) + \beta p_n(\sqrt{\lambda}) = 0$$
$$\alpha p_{n+1}(-\sqrt{\lambda}) + \beta p_n(-\sqrt{\lambda}) = 0 .$$

Ce qui montre que

$$p_{n+1}(\sqrt{\lambda})p_n(-\sqrt{\lambda}) - p_n(\sqrt{\lambda})p_{n+1}(-\sqrt{\lambda}) = 0$$

et que

$$p(t) = (\alpha p_{n+1}(t) + \beta p_n(t))/(t^2 - \lambda) .$$

D'autre part si $0 < u < \sqrt{\lambda}$, il est impossible que

$$p_{n+1}(u)p_n(-u) - p_n(u)p_{n+1}(-u) = 0 .$$

En effet, si cette dernière équation était réalisée, on pourrait trouver deux constantes α^* et β^* tels que $\alpha^* p_{n+1}(t) + \beta^* p_n(t) = (t^2 - u^2)p^*(t)$ et $(\alpha^*, \beta^*) \neq (0, 0)$. En posant $q^*(t) = (p^*(t))^2$, on aurait que

$$\int_0^\infty t^2(q^*(t))^2 e^{-t}dt = u^2 \int_0^\infty (q^*(t))^2 e^{-t}dt < \lambda \int_0^\infty (q^*(t))^2 e^{-t}dt .$$

Ce qui est contradictoire. C. Q. F. D.

Remarque. — θ_n est compris entre la plus petite racine de $p_{n+1}(t)$ et la plus petite racine de $p_n(t)$. On peut enfin transposer la démonstration du théorème 8 pour obtenir le résultat suivant.

THÉORÈME 11. — Si $L = \{ \lambda_1, \lambda_2, \ldots, \lambda_{2n-1} \}$ avec $0 < \lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_{2n-1}$ alors

$$\theta_n^2/\lambda_{2n-1}^2 \leq s(L) \leq \theta_n^2/\lambda_1^2 .$$

7. APPLICATION A L'APPROXIMATION PAR DES FONCTIONS DU TYPE DE BERNSTEIN

Considérons une suite $\{ \lambda_k \}_{k=1}^\infty$ où cette fois les λ_k ne forment plus une suite croissante mais que $0 < a \leq \lambda_k \leq b$ pour tout k. Soit $f(t)$ une fonction continue sur $[0, \infty)$ telle que $\lim_{t\to\infty} f(t)$ existe. Nous indiquons ici comment on peut approcher la fonction $f(t)$ de façon uniforme sur $[0, \infty)$ par un polynôme de Müntz dont les paramètres sont $\{ \lambda_k \}$. On se fixe une suite croissante d'entiers $\{ v_n \}$ telles que $\lim_{n\to\infty} n/v_n = \infty$. On choisit nv_n exposants de la suite $\{ \lambda_k \}$, on partage ces exposants en n blocs de v_n exposants $L_{1,n}, L_{2,n}, \ldots, L_{n,n}$. On considère alors les fonctions $f_{1,n}, f_{2,n}, \ldots, f_{n,n}$ telles que $f_{k,n} \in P^+(L_{k,n})$

$$\int_0^\infty f_{k,n}(t)dt = 1 \qquad \text{et} \qquad \int_0^\infty t^2 f_{k,n}(t)dt = s(L_{k,n}).$$

On introduit le tableau de variables aléatoires indépendantes entre elles

$$\{ \xi_{k,n} : 1 \leq k \leq n, n = 1, 2, \ldots \}$$

où la variable $\xi_{k,n}$ a une densité de probabilité égale à $f_{k,n}(t)$, $t \geq 0$.

Donnons en bref les raisons qui font que ce tableau se concentre totalement sur $[0, \infty)$. Vu le théorème 11, on peut trouver une constante universelle $C > 0$ telle que $s(L_{k,n}) \leq \dfrac{C}{a^2 v_n^2}$. L'inégalité de Schwarz donne que

$$\mathscr{E}(\xi_{k,n}) \leq (\mathscr{E}(\xi_{k,n}^2))^{\frac{1}{2}} \leq \frac{\sqrt{C}}{a v_n}.$$

Ainsi de façon uniforme en k, $\mathscr{E}(\xi_{k,n})$ tend vers zéro lorsque n tend vers l'infini. Le corollaire du théorème 8 donne que

$$\mathscr{E}(\xi_{k,n}) \geq m(L_{k,n}) \geq \frac{A}{a v_n}.$$

Ainsi

$$\sum_{k=1}^{m} \mathscr{E}(\xi_{k,n}) \geq \frac{Am}{a v_n}.$$

D'autre part

$$\sum_{k=1}^{m} \operatorname{Var}(\xi_{k,n}) \leq \sum_{k=1}^{m} \mathscr{E}(\xi_{k,n}^2) \leq \frac{Cm}{a^2 v_n^2}.$$

Ces inégalités suffisent pour vérifier la condition de concentration lorsque $n/v_n \to \infty$.

Si l'on pose $t_{k,n} = \mathscr{E}\left(\sum_{j=1}^{k} \xi_{j,n}\right)$ et si $X_n(t)$ est la suite de processus de naissance généralisés issus de ces deux tableaux, on aura que $f(t) = \lim_{n\to\infty} \mathscr{E}(f(X_n(t)))$ alors que la suite approchante est une somme d'exponentielles dont les paramètres viennent de la suite $\{ \lambda_k \}$ comme nous en assure le lemme 3.

On pourrait vouloir traiter le cas plus général d'une suite de nombres positifs λ_k qui admet 0 comme point d'accumulation. L. Schwartz [*8*] a montré par exemple que la condition nécessaire et suffisante pour que la famille de fonctions $\{ e^{-\lambda_k t} \}_{k=1}^{\infty}$ soit fondamentale dans $C_0[0, \infty)$ est que $\sum_{k=1}^{\infty} \lambda_k(1 + \lambda_k^2)^{-1} = \infty$. Il ne semble pas que notre méthode d'approximation d'une fonction f soit assez bonne pour obtenir ce résultat. Néanmoins celle-ci pourrait servir dans d'autres cadres, comme dans celui de l'approximation pondérée.

BIBLIOGRAPHIE

[1] S. BERNSTEIN, Démonstration du théorème de Weierstrass fondée sur le calcul des probabilités. *Commun. Soc. Math. Kharkow*, t. **13**, n° 2, 1912-1913, p. 1-2.

[2] A. O. GELFOND, On the generalized polynomials of S. N. BERNSTEIN (en russe). *Izv. Akad. Nauk SSSR*, ser. math., t. **14**, 1950, p. 413-420.

[3] I. I. HIRSCHMAN et D. V. WIDDER, Generalized Bernstein Polynomials. *Duke Math. J.*, t. **16**, 1949, p. 433-438.

[4] S. KARLIN et W. J. STUDDEN, *Tchebycheff Systems: with Applications in Analysis and Statistics*, Interscience, New York, 1966.

[5] G. G. LORENTZ, *Bernstein polynomials*. University of Toronto Press, Toronto, 1953.

[6] C. MUNTZ, *Uber den Approximationssatz von Weierstrass*. Mathematische Abhandlungen (Schwarzes Festschrift), Berlin, Springer, 1914, p. 303-312.

[7] E. PARZEN, *Stochastic Processes Holden-Day San Francisco*, 1962.

[8] L. SCHWARTZ, *Étude des sommes d'exponentielles réelles*. Hermann, Paris, 1943.

[9] G. SZEGO, Orthogonal Polynomials. *Amer. Math. Soc. Colloquium*. Publications, Vol. XXIII, Providence, 1939.

(Manuscrit reçu le 10 mars 1975)