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

I. H. DINWOODIE

Mesures dominantes et théorème de Sanov

Annales de l'I. H. P., section B, tome 28, nº 3 (1992), p. 365-373
<http://www.numdam.org/item?id=AIHPB_1992__28_3_365_0>

Ann. Inst. Henri Poincaré,
Vol. 28, n° 3, 1992, p. 365-373. *Probabilités et Statistiques*

Mesures dominantes et Théorème de Sanov

par

I. H. DINWOODIE
Tulane University, Department of Mathematics,
New Orleans, Louisiana 70118, U.S.A.

RÉSUMÉ. — On montre sous certaines conditions qu'une mesure de probabilité à la frontière de chaque boule ouverte de probabilités caractérise le taux de décroissance exponentielle de la probabilité que la loi empirique se trouve dans la boule. On en déduit des résultats plus fins que les théorèmes du type Sanov purement asymptotiques.

Mots clés : Grandes déviations, point dominant, théorème de Sanov.

ABSTRACT. — It is shown under certain conditions that for open balls of probability measures on a Polish space X, a unique probability measure determines the exponential decay rate of the probability that the empirical law L_n falls in the ball. Large deviation results stronger than purely asymptotic results can then be obtained.

1. INTRODUCTION

Soit X un espace topologique polonais avec distance d et soit $X_1, X_2, \ldots$ une suite de variables aléatoires, indépendantes et de même loi p, à valeurs dans X et définies dans (Ω, B, P). Soit L_n la loi empirique de l'échantillon

Classification A.M.S. : 60 F 10.

$\{X_i : 1 \leqq i \leqq n\}$:

$$\text{(1.1)} \qquad L_n(\omega, A) = \frac{1}{n} \sum_{i=1}^{n} 1_{\{X_i \in A\}}.$$

L'information de Kullback $\lambda(q)$ de la loi q par rapport à p détermine le comportement asymptotique de la loi aléatoire L_n sur le plan de grandes déviations. Plus précisément, soit **M** l'espace de mesures de probabilités sur X, muni de la topologie de la convergence étroite. Si $U \subset \mathbf{M}$ est un ouvert, et $C \subset \mathbf{M}$ est un fermé, on a

$$\text{(1.2)} \qquad \begin{cases} \liminf \frac{1}{n} \log \mathbf{P}(L_n \in U) \geqq -\Lambda(U) \\ \limsup \frac{1}{n} \log \mathbf{P}(L_n \in C) \leqq -\Lambda(C), \end{cases}$$

où $\Lambda(A) = \inf\{q \in A : \lambda(q)\}$, $A \subset \mathbf{M}$.

Théorèmes de ce genre remontent à Sanov [22], mais des résultats plus généraux se trouvent dans Bahadur et Zabell [3], Donsker et Varadhan [9], Groeneboom et Shorack [14], et Groeneboom, Oosterhoff, et Ruymgaart [15]. On s'intéresse ici à l'étude de (1.2) à l'aide de la notion d'un point dominant formulée par Ney ([18], [19]).

La notion d'un point dominant est à la base des fins résultats de Bahadur et Ranga Rao [2] sur les déviations de la moyenne $\bar{X}$ d'un échantillon de variables dans $\mathbf{R}^1$. L'idée se développe dans $\mathbf{R}^d$ par Ney. L'existence d'un point dominant en dimension finie, établi sous conditions par Ney, mène de la même façon à des améliorations utiles des résultats asymptotiques sur les probabilités de déviations de la moyenne $\bar{X}$. Ici on poursuit la formulation de Ney dans un espace de dimension infinie dans le but d'étudier le théorème de Sanov et donc le point dominant devient une mesure dominante. L'espace vectoriel topologique où se fait l'analyse est celui des mesures bornées sur X, muni d'une norme dont la topologie revient à la topologie de convergence étroite sur les mesures de probabilité.

Au fond, la technique excise le taux de décroissance exponentielle pour obtenir une intégrale relativement grande, a la portée d'un théorème limite central. En dimension finie, on a recours aux théorèmes locaux, comme ceux de Robinson *et al.* [21] ou de Stone [23], pour étudier l'intégrale. Mais en dimension finie on manque à présent de pareilles ressources.

Reeds [20] aborda l'étude raffinée de (1.2) pour la loi empirique dans le cas où l'espace X soit fini. L'analyse ici est dans un espace polonais, et il paraît que les problèmes géométriques qui s'élèvent dans les études de Reeds et Ney seront intéressants dans le cas général.

Applications statistiques des résultats comme (1.2) se trouvent dans la théorie de tests d'hypothèses et la théorie de la convergence d'estimateurs (*voir* respectivement Groeneboom et Shorack [14] ou Fortet et

Mourier [13]). Dans le cas où l'ensemble U ou C est le complémentaire d'un voisinage de la loi p de X_i, les limites à (1.2) représentent le taux de décroissance exponentielle du niveau de signification d'un test de l'hypothèse que X_i ait loi p (*voir* Manoukian [17] pour les définitions). Les limites (1.2) représentent aussi un taux de convergence de la loi empirique L_n vers la loi p.

Les notations suivantes seront employées. Soit E l'espace vectoriel de mesures bornées sur l'espace X. Soit $\mathbf{M} \subset E$ l'ensemble de mesures de probabilité, et soit L(X) l'espace de fonctions bornées et lipschitziennes définies dans X. Dans L(X), on a $L_{(a,b)}$, les fonctions qui sont lipschitziennes pour la constante $a>0$, et bornées par la constante $b>0$.

Définissons une norme $\|.\|$ sur l'espace de mesures E par

$$\|q\| = \sup_{f \in L_{(1,1)}} \left| \int f\,dq \right|.$$

La topologie du sous-espace **M** dans E est identique à la topologie de la convergence étroite sur **M** (Battacharya et Ranga Rao [4], p. 17). E′ désignera les formes linéaires et continues sur E.

Notons qu'un élément $f \in E'$ peut être regardé comme une fonction lipschitzienne et bornée lorsqu'elle est appliquée à un élément $q \in \mathbf{M}$. Plus précisément, si l'on définit une fonction f dans X par $f(x) = \langle f, \delta_x \rangle$, on a $\langle f, q \rangle = \int f\,dq$, pour chaque élément $q \in \mathbf{M}$. La fonction f est bornée, puisque $f(x) \leqq \sup_{\|q\| \leqq 1} |\langle f, q \rangle|$, et f est lipschitzienne :

$$\begin{aligned} |f(x) - f(y)| &\leqq \sup_{\|q\| \leqq 1} \langle f, q \rangle| \, \|\delta_x - \delta_y\| \\ &= \sup_{\|q\| \leqq 1} |\langle f, q \rangle| \sup_{g \in L_{(1,1)}} |g(x) - g(y)| \\ &\leqq \sup_{\|q\| \leqq 1} |\langle f, q \rangle| \, d(x, y). \end{aligned}$$

Du point de vue de la loi L_n, ce qui n'est pas dans **M** n'est pas important, et donc il est bien de considérer chaque élément $f \in E'$ comme une fonction bornée et lipschitzienne sur X.

Définissons la fonction $\lambda : E \to [0, \infty]$ par

$$\lambda(q) = \sup_{f \in E'} \left\{ \langle f, q \rangle - \log \int e^{f(x)} p(dx) \right\}.$$

Alors, selon le lemme suivant la fonction λ est tout simplement l'information de Kullback I, définie par

$$\begin{aligned} I(q) &= \int \log \left[\frac{dq}{dp} \right] dq \qquad && \text{si } q \ll p,\ q \in \mathrm{M}, \\ &= \infty && \text{autrement.} \end{aligned}$$

LEMME 1.1. — $I = \lambda$.

Preuve. — Il suffit de considérer le cas où $q \in \mathbf{M}$. Bahadur et Zabell ([3], p. 617) montrent que la borne supérieure à (1.2) en parcourant les fonctions continues et bornées est I. Il faut alors démontrer que la borne supérieure à (1.2) en parcourant l'espace $L(X)$ est supérieure à celle obtenue en parcourant les fonctions continues et bornées.

Soit $q \in \mathbf{M}$ une mesure fixe et soit g une fonction continue et bornée. D'après le théorème Stone-Weierstrass, à chaque compact $K \subset X$, il existe une suite de fonctions bornées et lipschitziennes qui converge uniformément vers f sur K. On en déduit que $\lambda(q) \geqq \langle g, q \rangle - \log \int e^{g(x)} p(dx)$, d'où le lemme. ■

L'information de Kullback $\lambda(q)$ est souvent désignée $K(q, p)$. La fonction λ est évidemment semi-continue inférieurement et convexe. De plus, le fermé $L_a = \{ q \in \mathbf{M} : \lambda(q) \leqq a \}$ est compact (Groeneboom *et al.* [15], Lemma 2.3).

Le symbole $\partial\lambda$ désignera le sous-différentiel de la fonction λ, et donc $\partial\lambda$ est contenu dans E', éventuellement vide. Soit $B \subset \mathbf{M}$. On appelle une mesure $q_B \in \mathbf{M}$ une *mesure dominante* de B si

$$(1.3) \quad \begin{cases} (a) \ q_B \in \bar{B}. \\ (b) \ \Lambda(B) = \lambda(q_B) \\ (c) \ \text{il existe } f \in \partial\lambda(q_B) \text{ telle que } f(q) \geqq f(q_B) \text{ pour toute mesure } q \in B. \end{cases}$$

Sanov [22] s'adressa au problème d'identifier une loi q qui satisfait la condition $\lambda(q) = \Lambda(B)$. Bolthausen [5] étudia aussi de tels points dans un espace de Banach, sans la condition (c). La condition (c) est importante dans la définition d'un point dominant dans $\mathbf{R}^d$ formulée par Ney [18], p. 159. Cette condition est à l'origine des difficultés techniques du développement ici.

Avant d'aborder la question de l'existence de mesures dominantes, nous rappelons la notion de I-projection généralisée de Csiszár [8] et son rapport avec une mesure dominante. La I-projection généralisée d'une mesure p sur un convexe $B \subset \mathbf{M}$ est la mesure p^* qui satisfait la condition suivante: chaque suite de probabilités $\{ q_n \} \subset B$ pour laquelle $K(q_n, p) \to K(B, p)$ converge vers p^* dans la topologie de variation totale, ou de valeur absolue de mesures. L'existence de la mesure p^* est établi par Csiszár [8], p. 769. La proposition suivante nous confirme qu'une mesure dominante d'un ouvert convexe n'est autre que la I-projection généralisée de la loi p.

PROPOSITION 1.1. — *Soit* $B \subset E$ *un ouvert convexe. Supposons que* $\Lambda(B) < \infty$. *Alors il existe une mesure unique* $q_B \in \bar{B}$ *telle que* $\lambda(q_B) = \Lambda(B)$, *et* q_B *est la* I-*projection généralisée de la mesure p sur* B.

Preuve. — Soit $\{ q_n \} \subset B$ une suite telle que $\lambda(q_n) \leqq \Lambda(B) + 1/n$. Le fermé $L_a = \{ q \in \mathbf{M} : \lambda(q) \leqq a \}$ est compact, et donc une suite extraite $\{ q_{n_k} \}$

converge vers une mesure $q_B \in \bar{B}$. On voit que $\lambda(q_B) \leqq \lim_k \lambda(q_{n_k}) = \Lambda(B)$, puisque λ est semicontinue inférieurement.

Mais la convexité de B entraîne aussi $\lambda(q_B) \geqq \Lambda(B)$. Pour tout réel $\alpha \in (0, 1)$ la mesure $\alpha q_B + (1-\alpha) q_1$ appartient aussi à B, et la convexité de λ entraîne

$$\lambda(q_B) \geqq \frac{1}{\alpha}[\lambda(\alpha q_B + (1-\alpha) q_1) - (1-\alpha)\lambda(q_1)] \geqq \frac{1}{\alpha}[\Lambda(B) - (1-\alpha)\lambda(q_1)].$$

Cette dernière quantité converge vers $\Lambda(B)$ lorsque $\alpha \to 1$, ce qui nous donne l'inégalité. Nous avons montré que $\lambda(q_B) = \Lambda(B)$ et aussi que $q_B \in \bar{B}$.

L'unicité de q_B est déduite des propriétés de convexité. On suppose d'abord qu'une mesure $m \in \bar{B}$ satisfait la condition $\lambda(m) = \Lambda(B)$. $\bar{B}$ est convexe et la mesure $\frac{1}{2}m + \frac{1}{2}q_B$ est en conséquence membre aussi de $\bar{B}$. La fonction $x \log(x)$ est strictement convexe, et on aurait l'inégalité

$$\lambda\left(\frac{1}{2}m + \frac{1}{2}q_B\right) < \Lambda(B) = \Lambda(\bar{B})$$

si les mesures m et q_B n'étaient pas identiques. Ce qui prouve que $m = q_B$, l'unicité de q_B.

Enfin, du fait que $\lambda(q_{n_k}) \to \Lambda(B)$, il en résulte la convergence de la suite extraite $\{q_{n_k}\}$ vers la I-projection généralisée p^* dans la topologie de variation totale. Mais $\{q_{n_k}\}$ converge vers q_B dans l'espace E, d'où $p^* = q_B$. ■

2. EXISTENCE

Le résultat principal est le théorème 2.1 qui établit l'existence d'une mesure dominante q_B et qui fournit une représentation de la probabilité $\mathbf{P}(L_n \in B)$ où figure une intégrale centrée par rapport à la loi q_B.

Définissons $M_n : X^n \to \mathbf{M}$ par

$$M_n(x_1, \ldots, x_n) = \frac{1}{n}\sum_{i=1}^{n} \delta_{x_i},$$

où δ_x est la mesure de Dirac au point $x \in X$.

THÉORÈME 2.1. – *Soit* $B \subset \mathbf{M}$ *un ouvert convexe, et on suppose que* $0 < \Lambda(B) < \infty$. *Alors il existe une mesure dominante* q_B *de* B *et on a*

$$(2.1) \qquad \mathbf{P}(L_n \in B) = e^{-n\Lambda(B)} \int_{\{M_n \in B\}} e^{-n\langle f, M_n - q_B\rangle} q_B^n(dx_1 \times \ldots \times dx_n),$$

où la forme f satisfait (1.3) *et*

$$e^{-\lambda(q_{\mathrm{B}})} q_{\mathrm{B}}(dx) = e^{f(x) - \langle f, q_{\mathrm{B}} \rangle} p(dx).$$

Preuve. — Soit q_{B} la mesure dans $\bar{\mathrm{B}}$ telle que $\lambda(q_{\mathrm{B}}) = \Lambda(\mathrm{B})$, dont l'existence est assurée par la Proposition 1.1. Il nous faut montrer la condition (1.3 *c*). Soient $g \in \mathrm{E}'$ une forme linéaire sur l'espace E et $\alpha \in \mathbf{R}$ un réel tels que

$$\langle g, q \rangle \leqq \alpha < \langle g, w \rangle$$

pour toute mesure q membre du convexe $\mathrm{L} = \{ v \in \mathrm{E} : \lambda(v) \leqq \Lambda(\mathrm{B}) \}$ et toute mesure w membre du convexe B (Bourbaki [6], p. II.40). Soit $\mathrm{U} = \{ q \in \mathrm{E} : \langle g, q \rangle > \alpha \}$. On déduit du Lemme 3.1 de Groeneboom *et al.* [15] que

$$\begin{aligned} -\Lambda(\mathrm{B}) &\leqq \liminf \frac{1}{n} \log \mathbf{P}(\langle g, \mathrm{L}_n \rangle > \alpha) \\ &\leqq \limsup \frac{1}{n} \log \mathbf{P}(\langle g, \mathrm{L}_n \rangle \geqq \alpha) \\ &\leqq -\Lambda(\{ q : \langle g, q \rangle \geqq \alpha \}) \\ &= -\Lambda(\bar{\mathrm{U}}) \\ &= -\Lambda(\mathrm{U}) \\ &= -\Lambda(\mathrm{B}); \end{aligned}$$

cette dernière égalité résulte de la convexité de l'ouvert U qui satisfait $\Lambda(\mathrm{U}) \leqq \Lambda(\mathrm{B}) < \infty$. Alors on a

$$\Lambda(\mathrm{B}) = -\lim \frac{1}{n} \log \mathbf{P}\left(\frac{1}{n} \sum_{i=1}^{n} g(\mathrm{X}_i) \geqq \alpha \right),$$

et en conséquence du théorème Cramér-Chernoff [1] et l'inégalité

$\mathrm{E}(g(\mathrm{X}_i)) = \langle g, p \rangle \leqq \alpha$, on a aussi

$$\begin{aligned} &= \sup_{t \in \mathbf{R}} t\alpha - \log \int e^{tg(x)} dp(x) \\ &= \sup_{t \in \mathbf{R}} t \langle g, q_{\mathrm{B}} \rangle - \log \int e^{tg(x)} dp(x) \\ &= t_0 \langle g, q_{\mathrm{B}} \rangle - \log \int e^{t_0 g(x)} dp(x), \end{aligned}$$

où le réel t_0 qui donne la borne supérieure existe parce que la variable $g(\mathrm{X}_i)$ est bornée. Mais cela revient à dire que $t_0 g \in \partial\lambda(q_{\mathrm{B}})$.

Il s'ensuit de l'inégalité $\mathrm{E}(g(\mathrm{X}_i)) = \langle g, p \rangle \leqq \alpha$ que $t_0 \geqq 0$. Ainsi la définition de la forme g donne pour toute $q \in \mathrm{L}$ et toute $w \in \mathrm{B}$

$$\langle t_0 g, q \rangle \leqq t_0 \alpha \leqq \langle t_0 g, w \rangle.$$

En particulier, puisque $q_{\mathrm{B}} \in \mathrm{L}$,

$$\langle t_0 g, w - q_{\mathrm{B}} \rangle \geqq 0$$

pour $w \in \mathrm{B}$, ce qui achève la démonstration de $(1.3c)$ lorsque on prend $f = t_0 g$.

On s'occupe maintenant de la représentation de $\mathbf{P}(\mathrm{L}_n \in \mathrm{B})$ à (2.1). A partir de la définition, nous avons

$$\begin{aligned}
\mathbf{P}(\mathrm{L}_n \in \mathrm{B}) &= \int_{\{\mathrm{M}_n \in \mathrm{B}\}} p^n(dx_1 \times \ldots \times dx_n) \\
&= e^{-n\lambda(q_{\mathrm{B}})} \int_{\{\mathrm{M}_n \in \mathrm{B}\}} \prod_{i=1}^{n} e^{-\langle f, \delta_{x_i} - q_{\mathrm{B}} \rangle} \mu^n(dx_1 \times \ldots \times dx_n) \\
&= e^{-n\lambda(q_{\mathrm{B}})} \int_{\{\mathrm{M}_n \in \mathrm{B}\}} e^{-n\langle f, \mathrm{M}_n - q_{\mathrm{B}} \rangle} \mu^n(dx_1 \times \ldots \times dx_n),
\end{aligned}$$

où μ désigne la probabilité sur X définie par

$$\mu(dx) = \frac{e^{f(x)}}{e^{\langle f, q_{\mathrm{B}} \rangle - \lambda(q_{\mathrm{B}})}} p(dx).$$

Il nous reste à établir que $\mu = q_{\mathrm{B}}$. Puisque la fonction λ est strictement convexe, il existe une mesure unique q_0 dans l'hyperplan

$$\mathrm{C} = \{ v \in \mathrm{E} : \langle f, v \rangle = \langle f, q_{\mathrm{B}} \rangle \}$$

qui satisfait la condition

$$\lambda(q_0) = \Lambda(\mathrm{C}) = \Lambda(\mathrm{B}) = \lambda(q_{\mathrm{B}}).$$

Il suffit donc de nous convaincre que $\mu \in \mathrm{C}$ et que $\lambda(\mu) = \lambda(q_{\mathrm{B}})$.

On a déjà l'identité

$$\langle f, q_{\mathrm{B}} \rangle - \log \int e^{f(x)} p(dx) = \sup_t \left\{ \langle tf, q_{\mathrm{B}} \rangle - \log \int e^{tf(x)} p(dx) \right\}.$$

La dérivée de la deuxième formule par rapport à t ainsi nous mène à l'égalité

$$\langle f, q_{\mathrm{B}} \rangle = \frac{\int f(x) e^{f(x)} p(dx)}{\int e^{f(x)} p(dx)} = \int f(x) \mu(dx) = \langle f, \mu \rangle,$$

en se servant du fait que $\int e^{f(x)} p(dx) = e^{\langle f, q_{\mathrm{B}} \rangle - \lambda(q_{\mathrm{B}})}$. Ça veut dire que $\mu \in \mathrm{C}$. Enfin, du lemme 1.1

$$\lambda(\mu) = \int (f(x) - \langle f, q_{\mathrm{B}} \rangle + \lambda(q_{\mathrm{B}})) \mu(dx) = \lambda(q_{\mathrm{B}}).$$

On a donc l'égalité $\mu = q_{\mathrm{B}}$, et ainsi s'achève la preuve. ■

On a immédiatement de la représentation (2.1) l'inégalité

$$\mathbf{P}(\mathrm{L}_n \in \mathrm{B}) \leqq e^{-n\Lambda(\mathrm{B})}$$

pour chaque $n \geqq 1$, qui vient de la condition (1.3 c). On peut faire mieux en combinaison avec les résultats de Bahadur et Ranga Rao [2] appliqués aux variables aléatoires $f(\mathrm{X}_i) \in \mathbf{R}$ et le demi-espace $\{q \in \mathrm{E} : \langle f, q \rangle \geqq \langle f, q_\mathrm{B} \rangle\}$ contenant B. On en déduit que

$$\mathbf{P}(\mathrm{L}_n \in \mathrm{B}) \leqq kn^{-1/2} e^{-n\Lambda(\mathrm{B})},$$

k étant une constante qui dépend de B. Fill [12] donne des résultats encore plus précis pour les demi-espaces. Mais le problème difficile, qu'on évite ici mais qui est à la base de toute utilisation sérieuse de (2.1), est l'étude de la géométrie de l'ensemble B au voisinage de q_B. Ce sera essentiel à la minoration de l'intégrale à (2.1) et à l'établissement des majorants plus précis, comme indiquèrent les études de Ney ([18], [19]) et Reeds [20] en dimension finie.

REMERCIEMENTS

Je tiens à remercier M. le referee de ses commentaires et John Liukkonen de quelques discussions du sujet.

RÉFÉRENCES

[1] R. AZENCOTT, *Grandes déviations et applications*, École d'Été de probabilités de Saint-Flour VIII-1978, LNM 774, Springer-Verlag, 1980, p. 1-176.

[2] R. R. BAHADUR et R. RANGA RAO, On Deviations of the Sample Mean, *Ann. Math. Stat.*, vol. **31**, 1960, p. 43-54.

[3] R. R. BAHADUR et S. L. ZABELL, Large Deviations of the Sample Mean in General Vector Spaces, *Ann. Probab.*, vol. **7**, 1979, p. 587-621.

[4] R. N. BATTACHARYA et R. RANGA RAO, *Normal Approximation and Asymptotic Expansions*, Wiley, 1976.

[5] E. BOLTHAUSEN, On the Probability of Large Deviations in Banach Spaces, *Ann. Probab.*, vol. **12**, 1984, p. 427-435.

[6] N. BOURBAKI, *Espaces Vectoriels Topologiques*, Masson, 1981.

[7] D. M. CHIBISOV, An Investigation of the Asymptotic Power of the Tests of Fit, *Theory Probab. Appl.*, vol. **10**, 1965, p. 421-437.

[8] I. CSISZÁR, Sanov Property, Generalized I-Projection and a Conditional Limit Theorem, *Ann. Probab.*, vol. **12**, 1984, p. 768-793.

[9] M. D. DONSKER et S. R. S. VARADHAN, Asymptotic Evaluation of Certain Markov Process Expectations for Large Time-III, *Commun. Pure Appl. Math.*, vol. **24**, 1976, p. 389-461.

[10] R. M. DUDLEY, Distances of Probability Measures and Random Variables, *Ann. Math. Stat.*, vol. **39**, 1968, p. 1563-1572.

[11] I. EKELAND et R. TEMAM, *Analyse convexe et problèmes variationnels*, Dunod, 1974.

[12] J. A. Fill, Asymptotic Expansions for Large Deviation Probabilities in the Strong Law of Large Numbers, *Probab. Theory Rel. Fields*, vol. **81**, 1989, p. 213-233.
[13] R. Fortet et E. Mourier, Convergence de la répartition empirique vers la répartition théorique, *Ann. Sci. École Norm. Sup.*, vol. **70**, 1953, p. 266-285.
[14] P. Groeneboom et G. P. Shorack, Large Deviations of Goodness-of-Fit Statistics and Linear Combinations of Order Statistics, *Ann. Probab.*, vol. **9**, 1981, p. 971-987.
[15] P. Groeneboom, J. Oosterhoff et F. H. Ruymgaart, Large Deviation Theorems for Empirical Probability Measures, *Ann. Probab.*, vol. **7**, 1979, p. 553-586.
[16] T. Inglot et T. Ledwina, On Probabilities of Excessive Deviations for Kolmogorov-Smirnov, Cramér-von Mises and Chi-Square Statistics, *Ann. Stat.*, vol. **18**, 1990, p. 1491-1495.
[17] E. B. Manoukian, *Modern Concepts and Theorems of Mathematical Statistics*, Springer, 1986.
[18] P. Ney, Dominating Points and the Asymptotics of Large Deviations for Random Walk on $\mathbf{R}^d$, *Ann. Probab.*, vol. **11**, 1983, p. 158-167.
[19] P. Ney, Convexity and Large Deviations, *Ann. Probab.*, vol. **12**, 1984, p. 903-906.
[20] J. Reeds, Correction Terms for Multinomial Large Deviations, *Asymptotic Theory of Statistical Tests and Estimation*, Academic Press, 1980, p. 287-305.
[21] J. Robinson, T. Höglund, L. Holst et M. P. Quine, On Approximating Probabilities for Small and Large Deviations in $\mathbf{R}^d$, *Ann. Probab.*, vol. **18**, 1990, p. 727-753.
[22] I. N. Sanov, On the Probabilities of Large Deviations of Random Variables, *Selected Translations in Mathematical Statistics and Probability* 1, American Mathematical Society, 1961, p. 213-244.
[23] C. Stone, On Local and Ratio Limit Theorems, *Proceedings of the Fifth Berkeley Symposium*, vol. **II**, University of California Press, Berkeley, 1967.

(*Manuscrit reçu le* 21 *mars* 1991;
révisé le 13 *janvier* 1992.)