C. R. Acad. Sci. Paris, Ser. I 334 (2002) 501–504

Probabilités/*Probability Theory*

Une inégalité de concentration à gauche pour les processus empiriques

Thierry Klein

Laboratoire de mathématiques, LAMA CNRS UMR 8100, Université de Versailles–Saint-Quentin, Versailles, France

Reçu le 31 janvier 2002 ; accepté le 4 février 2002

Note présentée par Michel Talagrand.

Résumé Nous donnons des constantes dans l'inégalité de concentration à gauche de Talagrand pour les processus empiriques indexés par des classes de fonctions, en partant de la méthode de Herbst. Le point nouveau est que la constante du facteur variance est exacte, ce qui répond à une conjecture de Massart. ***Pour citer cet article : T. Klein, C. R. Acad. Sci. Paris, Ser. I 334 (2002) 501–504.*** © 2002 Académie des sciences/Éditions scientifiques et médicales Elsevier SAS

Left concentration inequalities for empirical processes

Abstract We give new constants in Talagrand's left concentration inequality for maxima of empirical processes. Our approach is based on the Herbst method. The improvement we get concerns the constant in the variance factor, which is the one conjectured by Massart. ***To cite this article: T. Klein, C. R. Acad. Sci. Paris, Ser. I 334 (2002) 501–504.*** © 2002 Académie des sciences/Éditions scientifiques et médicales Elsevier SAS

1. Introduction

Soit $X_1, X_2, \ldots$ une suite de variables aléatoires indépendantes de loi commune P à valeur dans $\mathcal{X}$ espace polonais muni de sa tribu borélienne. Soit $\mathcal{F}$ une famille dénombrable de fonctions mesurables de $\mathcal{X}$ dans $\mathbb{R}$, de carré intégrable sous P. On pose

$$S_n(f) = f(X_1) + \cdots + f(X_n). \tag{1}$$

Nous regardons les propriétés de concentration de la variable aléatoire $Z = \sup\{S_n(f) : f \in \mathcal{F}\}$ autour de sa moyenne. Ce problème a été étudié dans une série de travaux successifs, en particulier par Talagrand [5] au moyen d'inégalités isopérimétriques pour les mesures produits. Il a obtenu une inégalité de type Bennett pour la déviation de Z par rapport à sa moyenne. Ensuite Ledoux [1] a montré que ces inégalités pouvaient aussi être obtenues à partir d'inégalités de type log-Sobolev pour les mesures produit. Cette technique permet de majorer la transformée de Laplace de Z. Massart [2] a clairement montré l'intérêt des méthodes entropiques (ou log-Sobolev) pour le calcul des constantes dans les inégalités de Talagrand, enfin Rio [4] a obtenu une inégalité de concentration à droite de type Bernstein avec un facteur variance asymptotiquement

Adresse e-mail : klein@math.uvsq.fr (T. Klein).

S1631-073X(02)02303-8/FLA

exact. Nous allons montrer dans cette note un analogue de l'inégalité de Rio [4] pour la concentration de Z à gauche de sa moyenne.

THÉORÈME 1.1. – *Soit $\mathcal{F}$ classe dénombrable de fonctions de $\mathcal{X}$ dans $]-\infty, 1]$, mesurables, de carré intégrable et d'espérance nulle sous P. Soit σ^2 telle que $\sigma^2 \geqslant P(f^2)$ pour toute f dans $\mathcal{F}$. Pour tout $x > 0$,*

- *si les fonctions de $\mathcal{F}$ sont à valeurs dans $[-1, 1]$ alors*

$$P\big(Z \leqslant E(Z) - x\big) \leqslant \exp\left(-\frac{v_n}{16} h\left(\frac{4x}{v_n}\right)\right). \tag{2}$$

Avec, $v_n = n\sigma^2 + 2E(Z)$ et $h(x) = (1+x)\log(1+x) - x$.

- *Si pour tout $f \in \mathcal{F}$, et tout $p \geqslant 2$, $|E(f(X_i)^p)| \leqslant \sigma^2 p!/2$ alors*

$$\forall t \in [0, 1[, \qquad L(t) \leqslant -t\,\mathrm{e}^{-t} E(Z) + \frac{n\sigma^2}{2} \frac{t^2(1+2t)}{1-t}, \tag{3}$$

$$P\big(Z \leqslant E(Z) - x\big) \leqslant \exp\left(-(u-v)^2\left(1 - 2\frac{v(u-v)}{u^2}\right)\right), \tag{4}$$

où $L(t)$ est la log-*Laplace de $-Z$, $u = \sqrt{x + v_n/2}$, $v = \sqrt{v_n/2}$.*

Remarque. – La fonction de taux de $v_n t^2(2(1-t))^{-1}$ est $(u-v)^2$, et nous avons en plus une correction multiplicative comprise ente $1/2$ et 1 tendant vers 1 quand $x v_n^{-1}$ tend vers 0.

2. Preuve

Comme dans Rio [4] on peut supposer que $\mathcal{F}$ est finie et procéder ensuite par passage à la limite. On définit $\mathcal{F}_n^k$ la tribu engendrée par $(X_i)_{i \neq k}$, $Z_k = \sup\{S_n^k(f) : f \in \mathcal{F}\}$. Puisque $\mathcal{F}$ est finie, on peut opérer une sélection $\mathcal{F}_n^k$-mesurable d'une fonction f_k telle que $Z_k = S_n^k(f_k)$, $Z_k' = S_n(f_k)$. On définit $\psi(x) = 1 - (1+x)\,\mathrm{e}^{-x}$, $\eta_k = f_k(X_k)$. Soit $t > 0$, alors le même calcul que dans Rio [4] nous conduit à l'inégalité suivante pour la transformée de Laplace F de $-Z$.

$$tF'(t) - F(t)\log F(t) \leqslant \sum_{k=1}^{n} E\big(\mathrm{e}^{-tZ_k}\psi\big(t(Z - Z_k)\big)\big). \tag{5}$$

On note alors que

$$\psi(tx) = \frac{t^2}{1-t} x\,\mathrm{e}^{-tx} + \left(1 - \mathrm{e}^{-tx}\left(\frac{1-t+tx}{1-t}\right)\right) = q(t)x\,\mathrm{e}^{-tx} + r(t,x), \tag{6}$$

la fonction $r(t,x)$ étant décroissante en x pour $0 \leqslant t < 1$ et $x \leqslant 1$. Comme $\sum_{k=1}^n (Z - Z_k) \leqslant Z$, et $q(t) > 0$, on déduit de (5) et (6) que

$$tF'(t) - F(t)\log F(t) \leqslant -q(t)F'(t) + \sum_{k=1}^{n} E\big(\mathrm{e}^{-tZ_k} r(t, Z - Z_k)\big). \tag{7}$$

Puisque $f_k(X_k) \leqslant Z - Z_k \leqslant 1$, et comme la fonction r est décroissante en x on a,

$$tF'(t) - F(t)\log F(t) \leqslant -q(t)F'(t) + \frac{1}{1-t}\sum_{k=1}^{n} E\big(\mathrm{e}^{-tZ_k}\big(1 - t - \mathrm{e}^{-t\eta_k}(1 - t + t\eta_k)\big)\big). \tag{8}$$

Posons $S_k = 1 - t - \mathrm{e}^{-t\eta_k}(1 - t + t\eta_k)$. L'équation (8) devient alors pour $0 \leqslant t < 1$

$$tF'(t) - (1-t)F(t)\log F(t) \leqslant \sum_{k=1}^{n} E\bigl(\mathrm{e}^{-tZ}\mathrm{e}^{t} E_n^k(S_k)\bigr), \tag{9}$$

où E_n^k est l'espérance conditionnelle par rapport à $\mathcal{F}_n^k$.

Sous la première hypothèse : on montre alors, à l'aide d'un développement en série entière de S_k que :

$$S_k \leqslant -t^2\eta_k + \eta_k^2\bigl(\mathrm{e}^t - 1 - t + t^3\bigr). \tag{10}$$

Par (2.4) dans Rio [4] on sait que $E_n^k(\eta_k) = 0$ et $E_n^k(\eta_k^2) \leqslant \sigma^2$. On obtient l'inéquation différentielle suivante valable pour tout $0 \leqslant t < 1$:

$$tL'(t) - (1-t)L(t) \leqslant n\,\mathrm{e}^t\sigma^2\bigl(\mathrm{e}^t - 1 - t + t^3\bigr). \tag{11}$$

On remarque que la fonction

$$G(t) = n\sigma^2\frac{\mathrm{e}^{4t} - 1 - 4t}{16} - t\,\mathrm{e}^{-t}E(Z), \tag{12}$$

est une sursolution de (11). On en déduit que pour tout $0 \leqslant t < 1$ on a :

$$P\bigl(Z \leqslant E(Z) - x\bigr) \leqslant \exp\left(v_n\frac{\mathrm{e}^{4t} - 1 - 4t}{16} - tx\right). \tag{13}$$

Le calcul usuel des grandes déviations donne alors (2) pour $x \leqslant v_n(\mathrm{e}^4 - 1)/4$. Pour obtenir (2) quand $x \geqslant v_n(\mathrm{e}^4 - 1)/4$ il suffit de remarquer que pour tout $f \in \mathcal{F}$ on a

$$P\bigl(Z \leqslant E(Z) - x\bigr) \leqslant P\bigl(S_n(f) \leqslant E(Z) - x\bigr), \tag{14}$$

et que le membre de droite de (14) peut être majoré par l'inégalité de Bennett classique par

$$P\bigl(S_n(f) \leqslant E(Z) - x\bigr) \leqslant \exp\left(-n\sigma^2 h\left(\frac{x - E(Z)}{n\sigma^2}\right)\right). \tag{15}$$

On conclut en remarquant que pour $x \geqslant v_n(\mathrm{e}^4 - 1)/4$, $x - E(Z) \geqslant 24x/25$, et ensuite que le majorant de (15) est à nouveau inférieur à la borne de (2).

Sous la deuxième hypothèse : comme $\eta_k = f_k(X_k)$ avec f_k $\mathcal{F}_n^k$-mesurable on a $|E_n^k(\eta_k^p)| \leqslant \sigma^2 p!/2$. On montre alors, à l'aide d'un développement en série entière de S_k que :

$$E_n^k(S_k) \leqslant \sigma^2\frac{t^2}{2(1-t)^2}\bigl(1 + 2t - 3t^2 + t^3\bigr). \tag{16}$$

On obtient l'inéquation différentielle suivante, valable pour tout $0 \leqslant t < 1$:

$$tL'(t) - (1-t)L(t) \leqslant \frac{n\sigma^2 t^2\,\mathrm{e}^t}{2(1-t)^2}\bigl(1 + 2t - 3t^2 + t^3\bigr). \tag{17}$$

Cette équation peut s'écrire

$$\left(\frac{L(t)}{t\,\mathrm{e}^{-t}}\right)' \leqslant \frac{n\sigma^2\,\mathrm{e}^{2t}}{2(1-t)^2}\bigl(1 + 2t - 3t^2 + t^3\bigr). \tag{18}$$

Définissons la fonction G par

$$G(t) = \frac{n\sigma^2\,\mathrm{e}^t(t + 2t^2)}{2(1-t)},$$

alors

$$G'(t) = \frac{n\sigma^2\,\mathrm{e}^t(1+5t-t^2-2t^3)}{2(1-t)^2}.$$

Nous allons montrer que pour tout t dans $I = [0,1[$

$$G'(t) \geqslant \frac{n\sigma^2\,\mathrm{e}^{2t}}{2(1-t)^2}\big(1+2t-3t^2+t^3\big).$$

Pour cela il suffit de montrer que

$$D(t) = 1+5t-t^2-2t^3-\mathrm{e}^t\big(1+2t-3t^2+t^3\big) \geqslant 0. \qquad (\mathcal{H})$$

On a

$$\begin{aligned} D^{(3)}(t) &= -12-\mathrm{e}^t\big(-5+2t+6t^2+t^3\big) \leqslant 0,\\ D^{(2)}(t) &= -2-12t-\mathrm{e}^t\big(-1-4t+3t^2+t^3\big), \quad D^{(2)}(0) = -1,\\ D'(t) &= 5-2t-6t^2-\mathrm{e}^t\big(3-4t+t^3\big). \end{aligned}$$

On en déduit que D' est décroissante, $(\mathcal{H})$ est alors une conséquence immédiate de la positivité de D aux bords de l'intervalle. Après majoration et intégration de (18) on obtient :

$$L(t) \leqslant -t\,\mathrm{e}^{-t}E(Z) + \frac{n\sigma^2}{2}\frac{t^2(1+2t)}{1-t} \leqslant -tE(Z) + \frac{v_n}{2}\frac{t^2(1+2t)}{1-t}. \qquad (19)$$

Par le calcul usuel de Cramer–Chernov on obtient

$$P\big(Z \leqslant E(Z)-x\big) \leqslant \exp\bigg(\frac{v_n}{2}\frac{t^2(1+2t)}{1-t} - tx\bigg). \qquad (20)$$

Pour obtenir une borne on prend alors le réel $t^\star$ minimisant $\frac{1}{2}v_n t^2(1-t)^{-1} - tx$. Après calcul on trouve $t^\star = 1 - w_n^{1/2}(x+w_n)^{-1/2}$, où $w_n = v_n/2$ que l'on réinjecte dans (20), pour obtenir (4). □

Remerciements. Je tiens à remercier E. Rio pour son aide précieuse.

Références bibliographiques

[1] M. Ledoux, On Talagrand's deviation inequalities for product measures, ESAIM Probability and Statistics 1 (1996) 63–87.

[2] P. Massart, About the constant in Talagrand's concentration inequalities for empirical processes, Ann. Probab. 28 (2000) 863–884.

[3] E. Rio, Théorie asymptotique des processus aléatoires faiblement dépendants, in: J.M. Ghidaglia, X. Guyon (Eds.), Math. Appl., Vol. 31, Springer, 2000.

[4] E. Rio, Une inégalité de Bernstein pour les maxima de processus empiriques, Preprint 57 de l'Université de Versailles–Saint-Quentin, 2001.

[5] M. Talagrand, New concentration inequalities in product spaces, Invent. Math. 126 (1996) 503–563.