Astérisque

DIDIER DACUNHA-CASTELLE

Formule de Chernoff pour une suite de variables réelles

Astérisque, tome 68 (1979), p. 19-24
<http://www.numdam.org/item?id=AST_1979__68__19_0>

Formule de Chernoff pour une suite de variables réelles

Didier DACUNHA-CASTELLE.

I - Le problème

Classiquement, on considère une suite $(X_n)_{n\in N}$ de variables aléatoires indépendantes équidistribuées, telles que :

$$\phi(t) = E \exp t X < \infty \quad \text{pour} \quad t \in]-M_1 , M_2[\quad , \quad M_1 , M_2 > 0 .$$

Ceci implique que EX existe, et que $S_n = X_1 + \ldots + X_n$ satisfait à $\frac{S_n}{n} \xrightarrow{p.s.} EX$.

Si l'on étudie les grandes déviations de $\frac{S_n}{n}$ par rapport à EX, à savoir la quantité $P(\frac{S_n}{n} > a)$, $a > EX$, (le cas $a < EX$ étant trivial), on a une majoration évidente qui est :

$$P(\frac{S_n}{n} > a) \leqq E \exp t (\frac{S_n}{n} - a) , \quad t > 0$$

ou encore
$$\frac{1}{n} \log P(S_n > na) \leqq \inf_{t>o} (\log \phi(t) - at)$$

à condition de convenir que $\phi(t) = \infty$, là où $\phi(t)$ n'est pas définie.

La formule de Chernoff consiste à obtenir une égalité à la place d'une majoration mais elle est simplement asymptotique. La fonction convexe $h(a) = \sup_{t>o} at - \log \phi(t)$ dont nous étudierons ci-dessous les propriétés, s'appelle le fonctionnelle de Cramer de (la loi de) X . Moyennant une définition correcte, elle permet de retrouver la loi de X tout comme ϕ . La passage de $\log \phi$ à h est une dualité classique sur les fonctions convexes dont nous rappellerons quelques points.

Enfin, il est essentiel de noter que la formule de Chernoff est conséquence de la seule existence des moments exponentiels et d'une certaine stationnarité. Pour cela, nous nous libérerons de l'hypothèse, S_n somme de variables indépendantes et équidistribuées pour la remplacer par : Z_n est une suite de variables aléatoires (d'abord réelles), telle que :

$$(1) \qquad \phi_n(t) = E \exp t\, Z_n < \infty \quad \text{pour} \quad t \in \left]-M_1, M_2\right[,\ M_1, M_2 > 0$$

et telle que :

$$(2) \qquad \frac{\log \phi_n(t)}{n} \longrightarrow L(t)\ , \quad (n \to \infty)\ , \quad t \in \left]-M_1, M_2\right[$$

soit $\phi_n^{1/n}(t) \to \exp L(t)$.

Cette dernière hypothèse impliquant évidemment que :

$$\frac{1}{n} E\, Z_n \longrightarrow L'(0)$$

(2) traduit la stationnarité minimale dans le problème et on cherchera à montrer que :

$$\lim_n \frac{1}{n} \log P(Z_n > na) = \inf_{t>o} (L(t) - at)$$

formule dite de Chernoff. Les étapes suivantes seront d'abord des études de cas particuliers et le passage à des variables à valeurs dans les espaces euclidiens ou fonctionnels, l'évènement $(Z_n > na)$ étant généralisé en $(\frac{Z}{n} \in G)$, G convenable.

II - Dualité et fonctions convexes

Soit f une fonction convexe sur R à valeurs dans $R \cup \{+ \infty\}$.

La duale $\hat{f}$ de f est définie par :

$$\hat{f}(t) = \sup_{x \in R} [tx - f(x)]$$

$\hat{f}$ est évidemment convexe, à valeurs dans $R \cup \{+ \infty\}$. Dans le cas où f est dérivable, alors $\hat{f}$ est dérivable, et $\hat{f}'$ est la fonction réciproque $(f')^{-1}$ de f'.

Les résultats de convergence suivants sont faciles (pour des fonctions de variables réelles). On note $Df = \{x, f(x) < \infty\}$:

1) Soit f_n une suite de fonctions convexes à valeurs dans $R \cup \{\infty\}$, convergeant vers f en tout point sauf peut-être les points de $(Df)^*$, frontière de Df. Supposons $D°f \neq \emptyset$, ($D°f$ intérieur de Df). Alors la convergence est uniforme sur tout compact de $D°f$ (et de $R - \overline{Df}$), ($\overline{Df}$ fermeture de Df).

2) De plus $\hat{f}_n(x) \longrightarrow \hat{f}(x)$ sauf poue les $x \varepsilon (D\hat{f})^*$.

3) Soit alors Z une variable aléatoire (non dégénérée).

$$\phi(t) = E \exp t Z$$

est convexe à valeurs dans $R \cup \{\infty\}$; $\log \varphi(t)$ est convexe (inégalité de Hölder).

Il est immédiat de vérifier que :

$$\{t , \log \phi(t) < \infty\} = \text{support } Z$$

et de manière plus précise, si $b = \sup(x , x \varepsilon \text{ support } Z\}$, on a :

$$\log \phi(b) < \infty \quad \text{équivaut à} \quad P(Z = b) = 0 .$$

Soit h la duale de $\log \phi$.

4) On a $h(EZ) = 0$, $h'(EZ) = 0$, h est croissante à droite de EZ.

5) Enfin, l'équation $(\log \varphi)'(t) = \frac{\phi'}{\phi}(t) = x_o$ a une solution si et seulement si $x_o \varepsilon \text{Support}° Z$ (ϕ étant dérivable sur $\text{Support}° Z\}$.

III - Formule de Chernoff

Soit Z_n une suite de variables aléatoires.

On suppose que Z_n vérifie les trois hypothèses suivantes :

H 1) $\quad E \exp t Z_n = \phi_n(t) < \infty \quad$ pour $\quad t \varepsilon]-M_1, M_2[$

$$0 < M_1 \ , \qquad 0 < M_2 \ .$$

H 2) $\phi_n^{1/n}(t) \longrightarrow \exp L(t)$, pour tout $t \in \,]-M_1, M_2[$

H 3) La suite Z_n n'est asymptotiquement dégénérée au sens où si C_n = support Z_n , $D_n = \frac{1}{n} C_n$ (homotéthique de C_n) , alors on suppose que

$\overset{\circ}{C_n} \neq \phi$, et $D = (\underline{\lim}_n D_n)^\circ \neq \phi$. (Dans le cas des variables indépendantes, équidistribuées X_i , H 1 et X_i non identiquement nulle implique H 3) .

Soit $B = \operatorname{int} L^{-1}(]M_1, M_2[)$.

On désigne par P_t^n Ia loi définie par $dP_t^n = \frac{e^{t} Z_n}{\phi_n(t)} dP$ et si f est mesurable, on note $E_t.f(Z_n)$ pour $E_{P_t^n} f(Z_n)$. Si $a \in D$, on sait d'après le § 2 , qu'il existe t_a^n unique tel que $\frac{1}{n} E_{t_a^n} Z_n = a$.

En général, $t_a^n \notin \,]M_1, M_2[$, cependant si $a \in B$, pour n assez grand $t_a^n \in \,]M_1, M_2[$ et :

$$\frac{1}{n} E_{t_a^n} Z_n = \frac{1}{n} \frac{\phi'_n(t_a^n)}{\phi_n(t_a^n)}$$

H 2) implique que $\frac{1}{n} \log \phi_n(t) \longrightarrow L(t)$ et que toutes les dérivées de $\frac{1}{n} \log \phi_n$ converge vers celles de L uniformément sur tout compact de $]-M_1, M_2[$.

Les duales h_n de $\frac{1}{n} \log \phi_n$ convergent simplement sur D, uniformément sur tout compact de D .

Remarquons enfin que l'on déduit de la croissance stricte de $\frac{1}{n} \frac{\phi'_n}{\phi_n}$ sur $]-M_1, M_2[$ et de la convergence de $\frac{1}{n} \frac{\phi'_n}{\phi_n}$ vers L' que $t_a^n \longrightarrow t_a$, $t_a \in \,]-M_1, M_2[$.

Si $t > t_a$, pour un η fixé , $\eta > 0$ assez petit , on a pour n assez grand , $\frac{1}{n} \frac{\phi'_n}{\phi_n}(t) > a-\eta$.

<u>Théorème</u>

Si Z_n vérifie H 1- 2- 3 , on a pour $a \in B$,

$\lim \frac{1}{n} \log P(Z_n > na) = -h(a)$, h duale de L .

Démonstration

$$P(Z_n > na) = \phi_n(t)\, E_t\, e^{-tZ_n}\, 1_{(Z_n > na)}$$

$$= \phi_n(t)\, e^{-nat}\, E_t\, e^{-nt(\frac{Z_n}{n} - a)}$$

$$\frac{1}{n} \log P(Z_n > na) > -at + \frac{1}{n} \log \phi_n(t)$$

$$+ \frac{1}{n} \log E_t\, e^{-nt(\frac{Z_n}{n} - a)}\, 1_{(\frac{Z_n}{n} - a > \varepsilon)}$$

(pour tout $\varepsilon > 0$)

$$> -at + \frac{1}{n} \log \phi_n(t)$$

$$+ \varepsilon t + \frac{1}{n} \log P_t\, (\frac{Z_n}{n} - a > \varepsilon)\ .$$

Comme $\lim \frac{Z_n}{n} > a$ P_t p.s. puisque $t > t_a$, on a $P_t(\frac{Z_n}{n} - a > \varepsilon) > \frac{1}{2}$ pour n assez grand, d'où le théorème, puisque ε est arbitraire.

Remarque : Nous avons utilisé la loi des grand nombres, sous les mêmes hypothèses en utilisant le théorème limite centrale, on obtient un peu plus. Reprenons en effet la démonstration en faisant $t = t_a^n$

$$\frac{1}{n} \log P(Z_n > na) > -a\, t_a^n + \frac{1}{n} \log \phi_n(t_a^n)$$

$$+ \frac{1}{n} \log E_{t_a^n} \exp \frac{Z_n - na}{\sqrt{n}\, \sigma_a} \sqrt{n}\, t_a^n \sigma_a\, 1_{(0 < \frac{Z_n - na}{\sqrt{n}\, \sigma_a} < \varepsilon)}$$

(ou $\sigma_a^2 = E_{t_a} (Y - a)^2 \neq 0$),

$$> - a t_a^n + \frac{1}{n} \log \phi_n(t_a^n) + \frac{\varepsilon t_a^n \sigma_a}{\sqrt{n}} + \frac{1}{n} \log H_1\ (]0,\varepsilon[)$$

ou $H_n\ ]0,\varepsilon[\ = P(0 < \frac{Z_n - na}{\sigma_a \sqrt{a}} < \varepsilon)$

$$\log E_{t_a^n} \exp t\left(\frac{Z_n - na}{\sqrt{n}\,\sigma_a}\right) = \log \phi_n\left(t_a^n + \frac{t}{\sigma_a \sqrt{a}}\right) - \log \phi_n(t_a^n) - \frac{at}{\sqrt{n}\,\sigma_a}$$

$$= \frac{t^2}{2n\,\sigma_a^2}\left(\frac{\phi'}{\phi}\right)'(t_a^n) + \frac{t^3}{6n^{3/2}\,\sigma_a^3}\left(\frac{\phi'}{\phi}\right)''\left(t_a + \alpha\frac{t}{\sigma_a\sqrt{n}}\right),$$

$0 < \alpha < 1$,

or $\frac{1}{n}\left(\frac{\phi'}{\phi}\right)'(t_a^n) \longrightarrow L''(t_a)$, $a \in B$

et $\left(\frac{\phi'}{\phi}\right)''\left(t_a + \alpha\frac{t}{\sigma\sqrt{n}}\right)$ est uniformément bornée (pour n assez grand)

par $2 \sup_{t\in K} |L''(t)|$, K compact de B incluant a .

Donc $\frac{Z_n - na}{\sqrt{n}\,\sigma_a} \longrightarrow$ loi normale sous $P_{t_a^n}$ et par suite $H_n\,]0,\varepsilon[\, > C(\varepsilon)$

pour n grand ou $C(\varepsilon) = \frac{1}{2} P(0 < G < \varepsilon)$, ou G est normale, d'où la

proposition. Sous les hypothèses du théorème, on a pour tout $\varepsilon > 0$,

$$\frac{1}{n}\log P(Z_n > na) > -h_n(t_a^n) + \frac{2\varepsilon\, t_a^n}{\sqrt{n}}\,\sigma_a^2 + \frac{C(\varepsilon)}{n}$$

en particulier dans le cas des variables indépendantes et équidistribuées

$$\frac{1}{n}\log P(Z_n > na) > -h(t_a) + \frac{\varepsilon t_a}{\sqrt{n}}\,\sigma_a^2 + \frac{2C(\varepsilon)}{n}$$

pour tout n grand.

<u>Remarque</u> : Au cours du chapitre suivant, nous aurons besoin du même théorème sous les hypothèses suivantes

$$\frac{1}{N} L_N(t) \to L(t) \text{ pour } t \in [0,1]$$

et $\frac{Z_N}{N}$ converge en P_0 probabilité.

Il est immédiat de modifier les démonstrations que nous venons de faire pour les appliquer à ce cas.

Didier DACUNHA-CASTELLE
Mathématiques - Bât. 425
ERA CNRS 532 "Statistique Appliquée"
Université Paris-Sud
91405 ORSAY.