Inégalité de concentration pour des variables aléatoires négativement associées

HAZAR ENNAFTI

Faculté des Sciences de Monastir

Journées ALEA 2011
07 Mars 2011

Inégalités de concentration pour des processus négativement associées

PLAN:

1. Variables aléatoires négativement associées
2. Inégalité de concentration
3. Application

Inégalités de concentration pour des processus négativement associées

PLAN:

1. Variables aléatoires négativement associées
2. Inégalité de concentration
3. Application

Inégalités de concentration pour des processus négativement associées

PLAN:

1. Variables aléatoires négativement associées
2. Inégalité de concentration
3. Application

Variables Aléatoires Négativement Associées

Définition

La famille des variables aléatoires indépendantes $\{X_1^*, ..., X_n^*\}$ est dite une copie indépendande de la famille des variables aléatoires $\{X_1, ..., X_n\}$ si X_i et X_i^* ont la même distribution et X_i est indépendante de X_i^* pour tout $i \in \{1, ..., n\}$.

Proposition

Soit $\{X_1, ..., X_n\}$ une famille des variables aléatoires et $\{X_1^*, ..., X_n^*\}$ une copie indépendante de $\{X_1, ..., X_n\}$ vérifiant $\forall i \in \{2, ..., n-1\}$, X_i^* est indépendante de $(X_1^*, ..., X_{i-1}^*, X_{i+1}, ..., X_n)$, on a pour toute fonction bornée $f : \mathbb{R}^n \to \mathbb{R}$ de classe C^n

$$Ef(X_1, ..., X_n) - Ef(X_1^*, ..., X_n^*) = \sum_{i=1}^{n-1} \int d_i(f)(t)\, dt. \tag{1}$$

avec $d_i(f)(t) = cov(\frac{\partial f}{\partial x_i}(X_1^*, ..., X_{i-1}^*, t, X_{i+1}, ..., X_n), \mathbf{1}_{X_i > t})$

Définition

Une famille des variables aléatoires $\{X_1, ...X_n\}$ est dite négativement associée (NA), si pour tous sous-ensembles disjoints A_1 et A_2 de {1,...,n} on a

$$cov(f_1(X_i, i \in A_1), f_2(X_j, j \in A_2)) \leq 0$$

où f_1 et f_2 sont deux fonctions croissantes par rapport à chaque composante.

Propriété

Deux fonctions croissantes, définies sur deux sous-ensembles disjoints des variables aléatoires négativement associées, sont négativement associées.

Définition

Une famille des variables aléatoires $\{X_1, ...X_n\}$ est dite négativement associée (NA), si pour tous sous-ensembles disjoints A_1 et A_2 de $\{$1,...,n$\}$ on a

$$cov(f_1(X_i, i \in A_1), f_2(X_j, j \in A_2)) \leq 0$$

où f_1 et f_2 sont deux fonctions croissantes par rapport à chaque composante.

Propriété

Deux fonctions croissantes, définies sur deux sous-ensembles disjoints des variables aléatoires négativement associées, sont négativement associées.

Corollaire

Soit $X_1, ..., X_n$ des variables aléatoires négativement associées et $\{X_1^*, ..., X_n^*\}$ une copie indépendante de $\{X_1, ..., X_n\}$ vérifiant $\forall i \in \{2, ..., n-1\}$, X_i^* est indépendante de $(X_1^*, ..., X_{i-1}^*, X_{i+1}, ..., X_n)$ alors

$$Ef(X_1, ..., X_n) - Ef(X_1^*, ..., X_n^*) \leq 0. \tag{2}$$

Pour toute fonction $f : \mathbb{R}^n \to \mathbb{R}$ vérifiant $t \mapsto \frac{\partial f}{\partial x_i}(x_1, ..., x_{i-1}, t, x_{i+1}, ..., x_n)$ est croissante $\forall i \in \{1, ..., n\}$.

Exemple

Soient $\{X_1, ..., X_n\}$ des variables aléatoires et $\{X_1^*, ..., X_n^*\}$ une cpoie indépendante de $\{X_1, ..., X_n\}$ vérifiant X_i^* est indépendante de $(X_1^*, ..., X_{i-1}^*, X_{i+1}, ..., X_n)$ pour tout $i \in \{1, ..., n\}$. On pose $f(X_1, ..., X_n) = g(\sum_{i=1}^{n} X_i)$, où g est une fonction convexe. Si $X_1, ..., X_n$ sont négativement associées, alors

$$cov(g'(\sum_{k=1}^{i-1} X_k^* + t + \sum_{k=i+1}^{n} X_k^*), \mathbf{1}_{X_i > t}) \leq 0$$

puisque g' et $\mathbf{1}_{X_i > t}$ sont croissantes pour tout $i = 1, ...n$. Ainsi

$$Eg(\sum_{i=1}^{n} X_i) - Eg(\sum_{i=1}^{n} X_i^*) \leq 0$$

Inégalité de Concentration

Théorème

Soit $g : R^n \to R$ une fonction bornée, on note $M = 2\|g\|_\infty$, vérifiant $\frac{\partial^2 g}{\partial x_i \partial x_j} + s\frac{\partial g}{\partial x_i}\frac{\partial g}{\partial x_j} \geq 0$ pour tout $s \geq 0$ et $j \geq i+1$. Soit $X_1, ..., X_n$ des variables aléatoires négativement associées et on pose $\{X_1^*, ..., X_n^*\}$ une cpoie indépendante de $\{X_1, ..., X_n\}$ tel que X_i^* est indépendante de $(X_1^*, ..., X_{i-1}^*, X_{i+1}, ..., X_n)$, alors on a $\forall t > 0$

$$P(g(X_1, ..., X_n) - Eg(X_1, ..., X_n) > t) \leq$$

$$\exp\{-\frac{varY_n}{M^2} h\left(\frac{M}{var(Y_n)}(t + D_n)\right)\} \tag{3}$$

avec $h(x) = (1+x)\log(1+x) - x$,
$Y_n = g(X_1^*, ..., X_n^*) - Eg(X_1^*, ..., X_n^*))$ et $D_n = \sum_{i=1}^{n-1} \int d_i(g)(t)\, dt$

Preuve:
En utilisant la majoration de *Chernoff*, on a pour tout $s > 0$

$$P(g(X_1, ..., X_n) - Eg(X_1, ..., X_n) > t) \leq$$

$$e^{-st} E(e^{s(g(X_1,...,X_n) - Eg(X_1,...,X_n))}) \tag{4}$$

On pose $f(x_1, ..., x_n) = \exp(s(g(x_1, ..., x_n) - Eg(X_1, ..., X_n)))$,
Alors on a $\forall j \geq i + 1 \qquad \frac{\partial^2 f}{\partial x_i \partial x_j}(x_1, ..., x_n) \geq 0$
Donc d'après l'inégalité (2) on a:

$$\begin{aligned} Ef(X_1, ..., X_n) &\leq Ef(X_1^*, ..., X_n^*) \\ &= E\left(\exp(s(g(X_1^*, ..., X_n^*) - Eg(X_1, ..., X_n)))\right) \end{aligned}$$

Reprenant le résultat (4)

$$P(g(X_1,...,X_n)-Eg(X_1,...,X_n)>t)\leq e^{-st}E(e^{s(g(X_1^*,...,X_n^*)-Eg(X_1,...,X_n))})$$
$$=e^{-st}e^{-sD_n}E(e^{sY_n})$$

avec $D_n=\sum_{i=1}^{n-1}\int d_i(g)(t)\,dt$ et $Y_n=g(X_1^*,...,X_n^*)-Eg(X_1^*,...,X_n^*))$. Notons que Y_n vérifie $EY_n=0$ et $\mid Y_n\mid\leq M=2\parallel g\parallel_\infty$, alors en utilisant l'astuce de *Bernstein* on a

$$E(e^{sY_n})=1+\sum_{k=2}^{\infty}s^k\frac{E(Y_n^k)}{k!}\leq exp(\sum_{k=2}^{\infty}\frac{s^kE(Y_n^k)}{k!})$$
$$=exp(\sum_{k=2}^{\infty}s^k\frac{var(Y_n)M^{k-2}}{k!})=exp(\frac{varY_n}{M^2}\sum_{k=2}^{\infty}s^k\frac{M^k}{k!})$$
$$=exp(\frac{varY_n}{M^2}(e^{sM}-1-sM))$$

Théorème

Et on a donc
$P(g(X_1,...,X_n) - Eg(X_1,...,X_n) > t)$
$\leq exp(-s(D_n + t + \frac{var(Y_n)}{M}) + \frac{var(Y_n)}{M^2} e^{sM} - \frac{var(Y_n)}{M^2})$
En optimisant le terme à droite, on obtient que le minimum est atteint pour

$$s_0 = \frac{1}{M} \log\left(\frac{M(t+D_n)}{var(Y_n)} + 1\right)$$

Ce qui donne alors

$$P(g(X_1,...,X_n) - Eg(X_1,...,X_n) > t) \leq \exp\{-\frac{varY_n}{M^2} h\left(\frac{M}{var(Y_n)}(t+D_n$$

avec $h(x) = (1+x)\log(1+x) - x$

Corollaire

Corollaire

Soit $X_1, ..., X_n$ des variables aléatoires négativement associées on a $\forall t > 0$

$$P(\sum_{i=1}^{n}(X_i - EX_i) > t) \leq \exp\{-\frac{\sum_{i=1}^{n}\sigma_i^2}{M^2} h\left(\frac{M(t + D_n)}{\sum_{i=1}^{n}\sigma_i^2}\right)\} \qquad (5)$$

avec $\sigma_i^2 = varX_i$, $M = 2\| \sum_{i=1}^{n} X_i\|_\infty$ et
$D_n = \sum_{i=1}^{n-1} \int \sum_{k=i+1}^{n} cov(X_k, \mathbf{1}_{X_i > t})\, dt$.

Remarques

1) Si les variables $X_1, ..., X_n$ sont indépendantes alors

$$D = 0 \; et \; Y_n \; = g(X_1, ..., X_n) - Eg(X_1, ..., X_n)$$

donc
$var(Y_n) = E\left(g(X_1, ..., X_n)\right)^2 - \left(Eg(X_1, ..., X_n)\right)^2 =$
$var(g(X_1, ..., X_n))$.
Si de plus $M = 1$, alors le résultat (3) donne

$$P(g(X_1, ..., X_n) - Eg(X_1, ..., X_n) > t) \leq$$

$$\exp\{-var(g(X_1, ..., X_n))h\left(\frac{t}{var(g(X_1, ..., X_n))}\right)\}$$

un résultat semblable à ce dernier apparait dans l'article de S. Boucheron, G. Lugosi et P. Massart , pour une fonction "Self bounding" et des variables aléatoires $X_1, ..., X_n$ indépendantes identiquement distribuées

$$P(g(X_1, ..., X_n) - Eg(X_1, ..., X_n) > t) \leq$$

$$\exp\{-Eg(X_1, ..., X_n) h\left(\frac{t}{Eg(X_1, ..., X_n)}\right)\}$$

où $h(x) = (1 + x)\log(1 + x) - x$.

Remarques

2) Dans le cas où les variables $X_1, ..., X_n$ sont réduites, indépendantes identiquement distribuées c'est à dire $D_n = 0$ et $\sigma_i^2 = 1 \ \forall i \in \{1, ..., n\}$, si $M = 1$ alors l'inégalité (5) devient

$$P(\sum_{i=1}^{n}(X_i - EX_i) > t) \leq \exp\{-nh\left(\frac{t}{n}\right)\}$$

$$= \exp\{-\left((n+t)\log(\frac{n+t}{n}) - t\right)\}$$

Application

Processus d'exclusion symetrique

On se donne d'une particule qui se déplace dans un ensemble dénombrable S selon les règles suivantes:

1. Il y a toujours une particule par site.
2. Une particule en x attend un temps aléatoire, puis elle se déplace vers le plus proche voision y avec une probabilité $p(x, y)$
$$(p(x, y) \geq 0 \; et \sum_y p(x, y) = 1)$$
3. Si y est vacant, alors elle se déplace vers y, et si y est occupée, elle reste à x.

Processus d'exclusion symetrique

Un processus η_t est dit *processus d'exclusion* correspondant à $p(x,y)$ lorsque $\eta_t(x) = 1$ si x est occupé et $\eta_t(x) = 0$ si x est vacant en un temps t.
Ainsi le générateur d'un processus d'exclusion η est donné par

$$Lf(\eta) = \sum_{\eta(x)=1,\eta(y)=0} p(x,y)(f(\eta_{x,y}) - f(\eta))$$

où $\eta_{x,y}$ est la configuration de η en interchangeant les coordonnées $\eta(x)$ et $\eta(y)$.
Le processus d'exclusion η est dit *symetrique* si $p(x,y) = p(y,x)$.

Définition

Soit μ une mesure de probabilité sur $\{0,1\}^n$ définie par $\mu = \mathcal{L}(\eta(1), ..., \eta((n)))$, on pose pour tous $z \in \mathbb{C}^n$ le polynôme générateur de μ

$$f(z) = f(z_1, ..., z_n) = E\prod_{i=1}^{n} z_i^{\eta(i)} \tag{6}$$

μ est dite "*Strong Rayleigh*" (*SR*) si pour tous $z \in \mathbb{R}$ on a

$$f(z)\frac{\partial^2 f}{\partial z_i \partial z_j}(z) \leq \frac{\partial f}{\partial z_i}(z)\frac{\partial f}{\partial z_j}(z) \tag{7}$$

μ est dite stable lorsque

$$f(z) \neq 0 si \quad Im(z_i) > 0 \quad \forall i \in \{1, ..., n\}$$

Proposition

Si la distribution initiale d'un processus d'exclusion symétrique est *SR* alors la distribution de η_t est négativement associée pour tous $t > 0$.

Preuve:
Sans perte de généralité, on considère le processus d'exclusion en deux cites, et on montre que la stabilité est préservée par la transformation

$$\mu \to T\mu = p\mu + (1-p)\mu_{ij}$$

où μ_{ij} est obtenu en permutant $\eta(i)$ et $\eta(j)$.
on suppose que f est stable et on montre, à l'aide du lemme suivant, que si $Im(z_k) > 0 \quad \forall k$ alors $Tf(z) \neq 0$.

Lemme: (Liggett)
soit $h(z,w) = a + bz + cw + dzw$ avec $a, b, c\, etd$ sont des complexes non nuls alors h est stable ssi

$$Re(b\bar{c} - a\bar{d}) > |bc - ad|$$

$Im(a\bar{d}) \geq 0$, $Im(a\bar{c}) \geq 0$, $Im(b\bar{d}) \geq 0$ et $Im(c\bar{d}) \geq 0$

D'où, si la distribution initiale d'un processus d'exclusion symetrique est stable alors la distribution en un temps $t > 0$ est stable. Ainsi, si la distribution initiale d'un processus d'exclusion symetrique est *SR* alors la distribution en un temps $t > 0$ est *SR* grace au résultat de P. Brandèn qui assure l'équivalence entre la propriété *SR* et la satabilité.
On retrouve, enfin, le résultat de la proposition à l'aide du théorème suivant:

Théorème:
Si μ est une mesure *SR* alors elle est négativement associée.

Conclusion

Si la distribution initiale d'un processus d'exclusion symétrique est *SR* alors la distribution η_t est négativement associée pour tous $t > 0$ et parsuite elle vérifie l'inégalité de concentration (3).

Bibliographie

- P. Brändén, *Polynomial with the half plane property and matroid theory*, Adv. Math 216 (2007), 302-320
- J. Borcea, P. Brändén and T. M. Liggett, *negative dependence and the geometry of polynomials 2008*
- S. Boucheron, G. Lugosi, P. Massart. A sharp concentration inequality with applications. Random Structures and Algorithms N°3 pp 288-292.
- K. Joag-Dev, F. Proschan. Negative association of random variables with application. Ann. Statist.11, 286-295.
- G. Lugosi. Concentration of measure inequalities. Febraury 9, 2006