Devoir en temps libre

à rendre avant 08.11.2013

Les inégalités de concentration décrivent la façon dont les variables aléatoires sont concentrées autour de leur moyenne ou médiane. Elles trouvent de nombreuses applications en probabilité, statistique, analyse, géométrie, algorithmique etc.. On se propose ici d'établir une telle inégalité à l'aide de la théorie des martingales puis on l'applique au problème du coloriage d'un graphe aléatoire et au problème isopérimétrique dans l'hypercube $\{0,1\}^n$.

Théorème 1 (Inégalité d'Azuma) *Soit $(X_n)_{n\geq 0}$ une martingale dont les accroissements sont contrôlés par une suite déterministe $(c_n)_{n\geq 1}$ i.e. presque sûrement $|X_n - X_{n-1}| \leq c_n$ pour tout $n \geq 1$. Alors, pour tout $\lambda > 0$, on a l'inégalité*

$$\mathbb{P}(|X_n - \mathbb{E}[X_n]| \geq \lambda) \leq 2e^{-\frac{\lambda^2}{2\sigma_n^2}}, \quad \textit{où} \quad \sigma_n^2 := \sum_{i=1}^n c_i^2.$$

Exercice 1 *Inégalité d'Azuma*

L'objet de ce premier exercice est d'établir le théorème ci-dessus. On considère donc une martingale $(X_n)_{n\geq 0}$ dont les accroissements sont contrôlés par une suite déterministe $(c_n)_{n\geq 1}$ i.e. presque sûrement $|X_n - X_{n-1}| \leq c_n$ pour tout $n \geq 1$.

1. Montrer que l'on peut se ramener au cas où $\mathbb{E}[X_n] = 0$ pour tout $n \geq 0$.
   La suite X_n étant une martingale, on a $\mathbb{E}[X_n] = \mathbb{E}[X_0]$ pour tout n, aussi il suffit de considérer la suite $X_n - \mathbb{E}[X_0]$ qui est centrée.
2. Montrer que si $u \in [-1, 1]$ et $t > 0$, on a les majorations $e^{ut} \leq \cosh(t) + u\sinh(t) \leq e^{t^2/2} + u\sinh(t)$.
   On écrit $ut = \frac{1+u}{2} \times t + \frac{1-u}{2} \times (-t)$. Par convexité de la fonction exponentielle, on a alors
   $$e^{ut} \leq \frac{1+u}{2} \times e^t + \frac{1-u}{2} \times e^{-t} = \cosh(t) + u\sinh(t) \leq e^{t^2/2}.$$
3. En déduire que si $t > 0$, on a la relation de récurrrence $\mathbb{E}[e^{tX_n}] \leq \exp\left(\frac{t^2c_n^2}{2}\right)\mathbb{E}[e^{tX_{n-1}}]$.
   On utilise la question précédente et la définition d'une martingale :
   $$\begin{aligned}\mathbb{E}\left[e^{tX_n}\right] &= \mathbb{E}\left[e^{t(X_n - X_{n-1})}e^{tX_{n-1}}\right] = \mathbb{E}\left[\mathbb{E}\left[e^{c_n t\frac{X_n - X_{n-1}}{c_n}} \mid \mathcal{F}_{n-1}\right] e^{tX_{n-1}}\right] \\ &\leq \cosh(c_n t)\mathbb{E}\left[e^{tX_{n-1}}\right] + \frac{\sinh(c_n t)}{c_n}\mathbb{E}\left[\mathbb{E}\left[X_n - X_{n-1} \mid \mathcal{F}_{n-1}\right] e^{tX_{n-1}}\right] \\ &\leq \exp\left(\frac{t^2c_n^2}{2}\right)\mathbb{E}\left[e^{tX_{n-1}}\right]\end{aligned}$$
   En itérant la relation de récurence, il vient
   $$\mathbb{E}\left[e^{tX_n}\right] \leq \exp\left(\frac{t^2\sum_{k=1}^n c_k^2}{2}\right).$$

4. Conclure en optimisant en $t > 0$.
Si $\lambda > 0$ et $t > 0$, par croissance de l'exponentielle et l'inégalité de Markov, on a
$$\begin{aligned}\mathbb{P}(X_n \geq \lambda) &= \mathbb{P}(e^{tX_n} \geq e^{\lambda t}) \leq e^{-\lambda t}\mathbb{E}\left[e^{tX_n}\right] \\ &\leq \exp\left(-\lambda t + \frac{t^2 \sum_{k=1}^n c_k^2}{2}\right).\end{aligned}$$
En optimisant en t, il vient $t^* = \lambda/\sigma_n^2$, soit
$$\mathbb{P}(X_n \geq \lambda) \leq e^{-\frac{\lambda^2}{2\sigma_n^2}}$$
On montre de même que
$$\mathbb{P}(X_n \leq -\lambda) \leq e^{-\frac{\lambda^2}{2\sigma_n^2}},$$
d'où le résultat.

Exercice 2 *Coloriage d'un graphe aléatoire à n sommets*
Étant donné un graphe G, le problème de son coloriage consiste à colorier les sommets de G avec la contrainte que deux sommets reliés par une arête doivent être de couleurs différentes. On appelle alors nombre chromatique de G le nombre minimal de couleurs nécessaires pour colorier le graphe. Par exemple, le nombre chromatique d'un graphe bipartite est 2, et celui d'un graphe planaire est au plus 4, c'est le fameux théorème des quatre couleurs. On s'intéresse ici au nombre chromatique d'un graphe aléatoire à n sommets qui possède donc au plus $N := \binom{n}{2}$ arêtes. Choisir un tel graphe au hasard revient précisément à décider, de façon aléatoire, pour chacune des N arêtes potentielles si elle appartient ou non au graphe en question. On considère ici le cas où la décision de garder ou non chaque arête est prise selon le résultat de variables aléatoires indépendantes $(Y_k)_{1\leq k\leq N}$ de loi Bernoulli de paramètre $0 < p < 1$. On considère ainsi l'espace de probabilités :
$$\Omega := \{0,1\}^{\binom{n}{2}}, \quad \mathcal{F} := \mathcal{P}(\Omega), \quad \mathbb{P}_p := \mathcal{B}(p)^{\otimes N}.$$
On introduit la filtration finie $(\mathcal{F}_k)_{0\leq k\leq N}$, où $\mathcal{F}_0 = \{\emptyset, \Omega\}$ et $\mathcal{F}_k := \sigma(Y_1, \ldots, Y_k)$ pour $1 \leq k \leq N$. Soit χ le nombre chromatique (aléatoire donc) du graphe aléatoire ainsi obtenu. On introduit la martingale $X_k := \mathbb{E}[\chi|\mathcal{F}_k]$, $0 \leq k \leq N$.

1. Montrer que la martingale (X_k) est à accroissements bornés.
Si deux graphes G_1 et G_2 coincident en dehors d'un ensemble d'arêtes qui contiennent un sommet commun v, alors on a $|\chi(G_1) - \chi(G_2)| \leq 1$ car on peut toujours attribuer une nouvelle couleur au sommet v et ainsi préserver la contrainte du coloriage.
2. En déduire que $\mathbb{P}_p(|\chi - \mathbb{E}[\chi]| \geq \lambda\sqrt{N}) \leq 2\exp\left(-\frac{\lambda^2}{2}\right)$.
La longueur de la filtration est N et on a $\mathcal{F}_N = \mathcal{F}$, le résultat est donc une application directe de l'inégalité d'Azuma.
3. En exhibant une autre filtration finie bien choisie, montrer que l'on a en fait
$$\mathbb{P}_p(|\chi - \mathbb{E}[\chi]| \geq \lambda\sqrt{n-1}) \leq 2\exp\left(-\frac{\lambda^2}{2}\right).$$
Vue la question précédente, il s'agit d'introduire une filtration de longeur $n-1$ c'est-à-dire une filtration $(\mathcal{G}_k)_{1\leq k\leq n-1}$ telle que $\mathcal{G}_1 = \{\emptyset, \Omega\}$ et $\mathcal{G}_{n-1} = \mathcal{F}$. Si $\{v_1, \ldots, v_n\}$ désigne la suite des n sommets du graphe, la tribu suivante convient :
$$\mathcal{G}_k := \sigma(\text{arête } e,\ e \subset \{v_1, \ldots, v_k\}), \quad 2 \leq k \leq n-1.$$

Exercice 3 *Inégalité isopérimétrique dans l'hypercube*
Dans le plan, l'inégalité isopérimétrique peut être formulée ainsi : à aire fixée, le disque est la surface dont le bord a le périmètre minimal. Une formulation équivalente et plus facilement généralisable à d'autres contextes est la suivante. Ici, d désigne la distance euclidienne, B_r la boule centrée en zéro de rayon $r > 0$; si $A \subset \mathbb{R}^2$, on note $|A|$ son aire et pour tout $\varepsilon > 0$, on définit l'$\varepsilon-$grossissement de A par $A_\varepsilon := \{x \in \mathbb{R}^2,\ d(x, A) \leq \varepsilon\}$.

Inégalité isopérimétrique : *Si $|A| = |B_r|$ alors, pour tout $\varepsilon > 0$, on a $|A_\varepsilon| \geq |B_{r+\varepsilon}|$.*

On s'intéresse ici au problème isopérimétrique lorsque l'espace sous-jacent n'est l'espace euclidien mais l'hypercube $\{0, 1\}^n$. L'aire euclidienne est maintenant remplacée par le cardinal et la distance d considérée est la distance de Hamming i.e. si $x, y \in \{0, 1\}^n$:

$$d(x, y) := \sum_{i=1}^{n} |x_i - y_i|.$$

En 1966, Harper a ainsi montré que si $|A| = |B_r|$ alors pour tout entier $k \geq 1$, on a $|A_k| \geq |B_{r+k}|$. Nous allons donner une preuve probabiliste d'une variante du résultat de Harper qui moralement signifie la chose suivante : si le cardinal d'un ensemble A représente une proportion non nulle ne celui de l'hypercube, alors son $\sqrt{n}-$grossissement contient presque tout l'hypercube.

Théorème 2 *Soient $A \subset \{0, 1\}^n$ et $\varepsilon > 0$ tels que $|A| \geq \varepsilon 2^n$. Soit $\lambda > 0$ tel que $\varepsilon = \exp(-\lambda^2/2)$. Alors, si $r = 2\lambda\sqrt{n}$, on a $|A_r| \geq (1 - \varepsilon)2^n$.*

Soit donc $A \subset \{0, 1\}^n$ et $\varepsilon > 0$ tels que $|A| \geq \varepsilon 2^n$. On munit $\Omega = \{0, 1\}^n$ de sa tribu des parties et de la mesure uniforme $\mathbb{P}$. Soit alors $Y = (Y_1, \ldots, Y_n)$ une variable de loi uniforme et X la distance de Hamming de Y à l'ensemble A. On introduit la filtration finie $(\mathcal{F}_k)_{0 \leq k \leq n}$ où $\mathcal{F}_0 = \{\emptyset, \Omega\}$ et $\mathcal{F}_k := \sigma(Y_1, \ldots, Y_k)$ pour $1 \leq k \leq n$. On introduit la martingale $X_k := \mathbb{E}[X|\mathcal{F}_k]$, $0 \leq k \leq n$.

1. Montrer que $|X_k - X_{k-1}| \leq 1$ pour $1 \leq k \leq n$.
   La fonction distance est $1-$lipschitzienne et lors du passage de X_{k-1} à X_k, seule la $k-$ième coordonnée est suceptible d'être modifiée.
2. À l'aide de l'inégalité d'Azuma (version unilatère), montrer que
   $$\mathbb{P}(X - \mathbb{E}[X] < -\lambda\sqrt{n}) \leq e^{-\frac{\lambda^2}{2}}, \tag{1}$$
   $$\mathbb{P}(X - \mathbb{E}[X] > \lambda\sqrt{n}) \leq e^{-\frac{\lambda^2}{2}}. \tag{2}$$
   Il s'agit d'une application directe de l'inégalité d'Azuma avec $(c_n) \equiv 1$.
3. De l'inégalité (1), déduire la majoration $\mathbb{E}[X] \leq \lambda\sqrt{n}$.
   D'après l'énoncé, on a $|A| \geq \varepsilon 2^n$ c'est-à-dire $\mathbb{P}(X = 0) \geq \varepsilon$. De l'inégalité (1), on déduit alors que $\mathbb{E}[X] \leq \lambda\sqrt{n}$.
4. De l'inégalité (2), déduire enfin que $\mathbb{P}(X \geq 2\lambda\sqrt{n}) \leq e^{-\frac{\lambda^2}{2}}$ et conclure.
   En injectant la dernière majoration dans l'inégalité (2), on déduit $\mathbb{P}(X \geq 2\lambda\sqrt{n}) \leq e^{-\frac{\lambda^2}{2}}$. En se rappelant que $\mathbb{P}$ est la loi uniforme, il vient
   $$\frac{|\{\omega \in \Omega,\ d(\omega, A) \geq 2\lambda\sqrt{n}\}|}{|\Omega|} \leq \varepsilon,$$
   ou de manière équivalente
   $$|\{\omega \in \Omega,\ d(\omega, A) \leq 2\lambda\sqrt{n}\}| \geq (1 - \varepsilon)2^n.$$