UNIVERSITÉ DE RENNES 1 — MASTER 1 MATHÉMATIQUE
MARTINGALES ET CHAÎNES DE MARKOV — ANNÉE 2014-2015

ÉLÉMENTS DE CORRECTION DU DM

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 issue de zéro 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*

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

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 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

$$e^{ut} \leq \frac{1+u}{2} \times e^t + \frac{1-u}{2} \times e^{-t} = \cosh(t) + u\sinh(t).$$

Puis, en developpant les fonctions $\cosh(t)$ et $e^{t^2/2}$ en série entière et en remarquant que $2n! \geq n!2^n$, on obtient facilement la majoration $\cosh(t) \leq e^{t^2/2}$.

2. 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 écarte le cas trivial où $c_n = 0$. Si $c_n > 0$, d'après la question précédente et en utilisant le fait que (X_n) est une martingale, on a :

$$\begin{aligned}\mathbb{E}\left[e^{tX_n}\right] &= \mathbb{E}\left[e^{t(X_n - X_{n-1})}e^{tX_{n-1}}\right] = \mathbb{E}\left[\mathbb{E}\left[e^{c_n t\frac{X_n - X_{n-1}}{c_n}} \mid \mathcal{F}_{n-1}\right]e^{tX_{n-1}}\right] \\ &\leq \cosh(c_n t)\mathbb{E}\left[e^{tX_{n-1}}\right] + \frac{\sinh(c_n t)}{c_n}\mathbb{E}\left[\mathbb{E}\left[X_n - X_{n-1} \mid \mathcal{F}_{n-1}\right]e^{tX_{n-1}}\right] \\ &\leq \exp\left(\frac{t^2c_n^2}{2}\right)\mathbb{E}\left[e^{tX_{n-1}}\right].\end{aligned}$$

En itérant la relation de récurrence, comme $X_0 = 0$, il vient

$$\mathbb{E}\left[e^{tX_n}\right] \leq \exp\left(\frac{t^2\sum_{k=1}^n c_k^2}{2}\right).$$

3. 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
$$\mathbb{P}(X_n \geq \lambda) \leq e^{-\frac{\lambda^2}{2\sigma_n^2}}$$
   La suite $(-X_n)$ vérifie les mêmes hypothèses que (X_n), on a donc aussi
$$\mathbb{P}(X_n \leq -\lambda) \leq e^{-\frac{\lambda^2}{2\sigma_n^2}},$$
   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 :
$$\Omega := \{0,1\}^N, \quad \mathcal{F} := \mathcal{P}(\Omega), \quad \mathbb{P}_p^{\otimes N} := \mathcal{B}(p)^{\otimes N}.$$
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.
   La donnée d'un graphe G à n sommets est équivalente à la donnée de la suite de ses arêtes, autrement dit si on énumère toutes les arêtes possibles entre 1 et N, à la donnée d'une suite $(y_1, \ldots, y_N) \in \{0,1\}^N$ qui est telle que $y_i = 1$ ssi la $i-$ème arête est effectivement présente dans le graphe considéré. Le nombre chromatique du graphe G est alors une fonction de la suite $(y_1, \ldots, y_N)$ considérée :
$$\chi(G) =: \psi(y_1, \ldots, y_N).$$
   On remarque que si deux suites diffèrent d'exactement un facteur, alors les nombres chromatiques correspondants diffèrent d'au plus d'un, i.e. si $z \in \{0,1\}$:
$$|\psi(y_1, \ldots, y_k, z, y_{k+2}, \ldots, y_N) - \psi(y_1, \ldots, y_k, 1-z, y_{k+2}, \ldots, y_N)| \leq 1.$$
   En effet, si la "nouvelle arête" pose un problème pour le coloriage, il suffit de colorier un sommet à l'une de ses extrémités avec une nouvelle couleur pour obtenir un coloriage licite.

On choisit ici la suite $(y_1, \ldots, y_N)$ comme réalisation d'une variable aléatoire $(Y_1, \ldots, Y_N)$ à valeurs $\{0,1\}^N$ de loi $\mathbb{P}_p$, obtenant ainsi un graphe aléatoire G à n sommets. Par définition, la suite X_n est donnée par

$$X_n = \mathbb{E}[\chi(G)|\mathcal{F}_n] = \mathbb{E}[\psi(Y_1, \ldots, Y_N)|Y_1, \ldots, Y_n].$$

Les variables Y_i étant indépendantes et de même loi, on a

$$X_n = \int_{\{0,1\}^{N-n}} \psi\left(Y_1, \ldots, Y_n, z_{n+1}, \ldots, z_N\right) d\mathbb{P}_p(z_{n+1}) \times \ldots \times d\mathbb{P}_p(z_N).$$

et de même

$$X_{n+1} = \int_{\{0,1\}^{N-n-1}} \psi\left(Y_1, \ldots, Y_n, Y_{n+1}, z_{n+2}, \ldots, z_N\right) d\mathbb{P}_p(z_{n+2}) \times \ldots \times d\mathbb{P}_p(z_N).$$

Comme $Y_{n+1} = Y_{n+1}(\omega)$ ne dépend pas de z_{n+1}, on peut également écrire X_{n+1} de façon artificielle comme une intégrale sur l'espace $\{0,1\}^{N-n}$:

$$X_{n+1} = \int_{\{0,1\}^{N-n}} \psi\left(Y_1, \ldots, Y_n, Y_{n+1}, z_{n+2}, \ldots, z_N\right) d\mathbb{P}_p(z_{n+1}) \times d\mathbb{P}_p(z_{n+2}) \times \ldots \times d\mathbb{P}_p(z_N).$$

Auquel cas l'accroissement $X_{n+1} - X_n$ s'exprime aussi comme une intégrale

$$X_{n+1} - X_n = \int_{\{0,1\}^{N-n}} A \, d\mathbb{P}_p(z_{n+1}) \times \ldots \times d\mathbb{P}_p(z_N),$$

où l'intégrande $A = A(\omega)$ est naturellement donnée par la différence

$$A := \psi\left(Y_1, \ldots, Y_n, z_{n+1}, z_{n+2}, \ldots, z_N\right) - \psi\left(Y_1, \ldots, Y_n, Y_{n+1}, z_{n+2}, \ldots, z_N\right).$$

D'après la remarque ci-dessus, on a $|A| \leq 1$ presque sûrement car les deux suites intervenant ici ne diffèrent au plus que par la $(n+1)-$ième coordonnée. On a donc presque sûrement

$$|X_{n+1} - X_n| \leq \int_{\{0,1\}^{N-n}} |A| \, d\mathbb{P}_p(z_{n+1}) \times \ldots \times d\mathbb{P}_p(z_N) \leq 1.$$

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 tribu $\mathcal{F}_0$ est triviale donc $X_0 = \mathbb{E}[\chi]$, par ailleurs $\mathcal{F}_N = \mathcal{F}$ donc $X_N = \chi$. En appliquant l'inégalité d'Azuma à la suite $(X_n - X_0)$ issue de zéro et dont les accroissements sont ceux de la suite (X_n), on obtient le résultat voulu.
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 :

$$\mathcal{G}_k := \sigma(\text{arête } e, \ e \text{ a une extrémité dans } \subset \{v_1, \ldots, v_k\}), \quad 2 \leq k \leq n-1.$$

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 $1 \leq k \leq n$.
   On raisonne comme dans la première question de l'exercice 2. Pour simplifier les expressions, on note d_A la distance à A et $\mathbb{P}_{1/2}$ la loi uniforme sur $\{0,1\}$. On a alors
   $$X_{k+1} - X_k = \int_{\{0,1\}^{n-k}} B \, d\mathbb{P}_{1/2}(z_{k+1}) \times \ldots \times d\mathbb{P}_{1/2}(z_n),$$
   où l'intégrande $B = B(\omega)$ est donnée par la différence
   $$B := d_A(Y_1, \ldots, Y_k, z_{k+1}, z_{k+2}, \ldots, z_n) - d_A(Y_1, \ldots, Y_k, Y_{k+1}, z_{k+2}, \ldots, z_n).$$
   Le point crucial est que la fonction distance est $1-$lipschitzienne de sorte que l'intégrande vérifie $|B| \leq 1$. En intégrant, on obtient ainsi le contrôle $|X_{k+1} - X_k| \leq 1$ presque sûrement.
2. À l'aide de l'inégalité d'Azuma (version unilatère), montrer que
   $$\mathbb{P}(X - \mathbb{E}[X] < -\lambda\sqrt{n}) \leq e^{-\frac{\lambda^2}{2}}, \tag{1}$$
   $$\mathbb{P}(X - \mathbb{E}[X] > \lambda\sqrt{n}) \leq e^{-\frac{\lambda^2}{2}}. \tag{2}$$
   Comme dans l'exercice 2, il s'agit d'une application directe de l'inégalité d'Azuma avec $c_n \equiv 1$. Notez que les majorations sont vérifiées pour tout $\lambda > 0$ (pas seulement pour le λ associé à ε). Par ailleurs, les inégalités strictes dans les probabilités découlent des inclusions
   $$\{X - \mathbb{E}[X] < -\lambda\sqrt{n}\} \subset \{X - \mathbb{E}[X] \leq -\lambda\sqrt{n}\}, \quad \{X - \mathbb{E}[X] > \lambda\sqrt{n}\} \subset \{X - \mathbb{E}[X] \geq \lambda\sqrt{n}\}.$$

3. De l'inégalité (1), déduire la majoration $\mathbb{E}[X] \leq \lambda\sqrt{n}$.

   Soit $\lambda' > \lambda$ de sorte que $\varepsilon' := \exp(-{\lambda'}^2/2) < \varepsilon$. D'après l'énoncé, on a $|A| \geq \varepsilon 2^n$ c'est-à-dire $\mathbb{P}(X = 0) \geq \varepsilon$ et en particulier $\mathbb{P}(X = 0) > \varepsilon'$. Supposons par l'absurde l'inégalité $\mathbb{E}[X] > \lambda'\sqrt{n}$. Alors, de l'inégalité (1), on déduit

$$\begin{aligned}\mathbb{P}(X = 0) \quad &= \mathbb{P}(\{X - \mathbb{E}[X] < -\lambda'\sqrt{n}\} \cap \{X = 0\}) \\ &\leq \mathbb{P}(X - \mathbb{E}[X] < -\lambda'\sqrt{n}) \\ &\leq e^{-\frac{{\lambda'}^2}{2}} \\ &\leq \varepsilon',\end{aligned}$$

   ce qui vient contredire $\mathbb{P}(X = 0) > \varepsilon'$. Nécessairement, on a donc $\mathbb{E}[X] \leq \lambda'\sqrt{n}$ et ce pour tout $\lambda' > \lambda$, d'où le résultat.

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

$$|\{\omega \in \Omega,\, d(\omega, A) \leq 2\lambda\sqrt{n}\}| \geq (1 - \varepsilon)2^n.$$