Apprentissage statistique: TD1
Inégalités de concentration

Emile Contal
http://econtal.perso.math.cnrs.fr/teaching

16 janvier 2015

Exercice 1.

1. **(Inégalité de Markov)** Soit X une variable aléatoire réelle positive. Montrer l'inégalité suivante :
$$\forall a > 0,\ \ \Pr[X \geq a] \leq \frac{\mathbf{E}[X]}{a}\,.$$
2. **(Corollaire)** On dit qu'une variable aléatoire réelle X est *b-sous-Gaussienne* lorsque pour tout $s \in \mathbb{R}$ on a $\mathbf{E}[\exp(sX)] \leq \exp(\frac{s^2b^2}{2})$. Montrer alors :
$$\forall a > 0,\ \ \Pr[X \geq a] \leq \exp\Big(-\frac{a^2}{2b^2}\Big)\,.$$
3. **(Propriété)** Montrer que lorsque X est b-sous-Gaussienne, alors $\mathbf{E}[X] = 0$ et $\mathbf{Var}[X] \leq b^2$.

Exercice 2.

1. **(Somme de variables indépendantes)** Soit $X_1, \dots, X_n$ une famille de variables aléatoires réelles indépendantes telles que pour tout $i \leq n,\ X_i$ est b_i-sous-Gaussienne. Soit $S_n = X_1 + \dots + X_n$. Montrer que :
$$\Pr[S_n \geq a] \leq \exp\Big(-\frac{a^2}{2\sum_{i=1}^n b_i^2}\Big)\,.$$
2. **(cas Gaussien)** Soit X une variable aléatoire Gaussienne $X \sim \mathcal{N}(\mu, \sigma^2)$. Montrer que $X - \mathbf{E}[X]$ est σ-sous-Gaussienne. En déduire une inégalité de concentration sur la somme de n Gaussiennes indépendantes.
3. **(cas Rademacher)** Soit X une variable aléatoire suivant une loi de Rademacher, c'est à dire $\Pr[X = -1] = \Pr[X = 1] = \frac{1}{2}$. Montrer que X est 1-sous-Gaussienne. En déduire une inégalité de concentration sur la somme de n variables de Rademacher indépendantes.
Indice : utiliser $2^n n! \leq (2n)!$.

4. **(Inégalité de Hoeffding)** Soit X une variable aléatoire réelle telle que $X \in [c,d]$. Montrer que $X - \mathbf{E}[X]$ est sous-Gaussienne de paramètre $\frac{d-c}{2}$. En déduire une inégalité de concentration sur la somme de n variables aléatoires indépendantes bornées.
Indices : utiliser la convexité de l'exponentielle pour la borner par une corde entre c et d ; écrire la borne obtenue sous la forme $e^{L(h,p)}$ *avec* $h = s(d-c)$ *et* $p = \frac{-c}{d-c}$ *; borner* $L(h,p)$ *grâce à une décomposition de Taylor au second ordre en h.*

Exercice 3.

1. **(Espérance conditionnelle)** Soit $(\Omega, \mathcal{F}, P)$ un espace de probabilité, $\mathcal{F}'$ une sous-tribu de $\mathcal{F}$ et X une variable aléatoire dans $L^1(\Omega, \mathcal{F}, P)$. On définit l'*espérance conditionnelle de* X *sachant* $\mathcal{F}'$ notée $\mathbf{E}[X \,|\, \mathcal{F}']$ comme étant l'unique variable aléatoire dans $L^1(\Omega, \mathcal{F}', P)$ telle que pour toute variable aléatoire Z $\mathcal{F}'$-mesurable bornée :
$$\mathbf{E}[ZX] = \mathbf{E}[Z\mathbf{E}[X \,|\, \mathcal{F}']]\,.$$
On a également, $\forall B \in \mathcal{F}',\ \mathbf{E}[X\mathbb{1}_B] = \mathbf{E}[\mathbb{1}_B\mathbf{E}[X \,|\, \mathcal{F}']]$.
Lorsque $\mathcal{F}'$ est une tribu engendrée par une variable aléatoire Y, c'est à dire $\mathcal{F}' = \sigma(Y)$, on utilisera la notation $\mathbf{E}[X \,|\, Y]$.
Montrer les propriétés suivantes :
   (a) L'application $X \to \mathbf{E}[X \,|\, \mathcal{F}']$ est linéaire.
   (b) $\mathbf{E}[\mathbf{E}[X \,|\, \mathcal{F}']] = \mathbf{E}[X]$.
   (c) Si X est $\mathcal{F}'$-mesurable, $\mathbf{E}[X \,|\, \mathcal{F}'] = X$.
   (d) Si X et $\mathcal{F}'$ sont indépendants, $\mathbf{E}[X \,|\, \mathcal{F}'] = \mathbf{E}[X]$.
   (e) Si $\mathcal{F}'' \subseteq \mathcal{F}' \subseteq \mathcal{F}$, alors $\mathbf{E}[\mathbf{E}[X \,|\, \mathcal{F}''] \,|\, \mathcal{F}'] = \mathbf{E}[\mathbf{E}[X \,|\, \mathcal{F}'] \,|\, \mathcal{F}''] = \mathbf{E}[X \,|\, \mathcal{F}'']$.
2. **(Martingales)** Soient $(X_n)_{n\geq 0}$ une famille de variables aléatoires indexée par $\mathbb{N}$ et $(\mathcal{F}_n)_{n\geq 0}$ une suite de tribus croissante pour l'inclusion, c'est à dire une *filtration*. On dit que $(X_n)_{n\geq 0}$ est *adapté* à $(\mathcal{F}_n)_{n\geq 0}$ lorsque X_n est $\mathcal{F}_n$-mesurable pour tout n. De plus si pour tout n on a X_n intégrable et :
$$\mathbf{E}[X_{n+1} \,|\, \mathcal{F}_n] = X_n\,,$$
alors $(X_n)_{n\geq 0}$ est une *martingale*. Lorsque $\mathcal{F}_n = \sigma(X_0, \ldots, X_n)$, on parle de la *filtration naturelle* de $(X_n)_{n\geq 0}$.
Soit $(X_n)_{n\geq 0}$ une martingale adaptée à $(\mathcal{F}_n)_{n\geq 0}$ telle que pour tout $s \in \mathbb{R}$ on ait $\mathbf{E}[\exp\big(s(X_{n+1} - X_n)\big) \,|\, \mathcal{F}_n] \leq \exp(\frac{s^2 b_{n+1}^2}{2})$. Montrer que pour tout $a > 0$:
$$\Pr[X_n - X_0 \geq a] \leq \exp\Big(-\frac{a^2}{2\sum_{i=1}^n b_i^2}\Big)\,.$$
3. **(Inégalité d'Azuma)** En déduire une inégalité de concentration pour une martingale $(X_n)_{n\geq 0}$ telle que $|X_n - X_{n-1}| \leq c_n$.