UPS TOULOUSE III **MASTER 2**
2005/2006

TD I
ALGORITHMES STOCHASTIQUES

1 Robbins-Monro.

Soit f une densité de probabilité, paire et continue sur $\mathbb{R}$. On observe un modèle de translation

$$X_n = \theta + Z_n$$

où (Z_n) est une suite de variables aléatoires réelles indépendantes et de même loi, de densité de probabilité f et où θ est un paramètre réel inconnu à estimer. Par analogie avec la procédure de Robbins-Monro, on estime θ par l'estimateur récursif

$$\widehat{\theta}_{n+1} = \widehat{\theta}_n - \gamma_n \Big(\mathbb{1}_{(X_{n+1} \leq \widehat{\theta}_n)} - \frac{1}{2} \Big)$$

avec $\widehat{\theta}_0$ choisi arbitrairement et où la suite (γ_n) est positive, décroissante vers zéro avec

$$\sum_{n=1}^{\infty} \gamma_n = +\infty \qquad \text{et} \qquad \sum_{n=1}^{\infty} \gamma_n^2 < +\infty.$$

1) Montrer que $\widehat{\theta}_n \longrightarrow \theta$ p.s.

2) Dans le cas particulier où f est strictement positive et $\gamma_n = 1/(nf(0))$, montrer le théorème limite centrale

$$\sqrt{n}(\widehat{\theta}_n - \theta) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \sigma^2),$$

avec $\sigma^2 > 0$ à déterminer.

2 Pas.

Dans l'algorithme de Robbins-Monro pour le dosage

$$X_{n+1} = X_n + \gamma_n (Y_{n+1} - \alpha),$$

la suite positive et décroissante vers zéro (γ_n) doit satisfaire à la condition

$$\sum_{n=0}^{\infty} \gamma_n = +\infty.$$

Si cette condition n'est pas satisfaite, (X_n) ne converge pas nécessairement vers x^*. Soit (ε_n) une suite de variables aléatoires indépendantes et de même loi $\mathcal{N}(0, 1)$. On suppose que

$$Y_{n+1} = \theta X_n + \varepsilon_{n+1} \qquad \text{et} \qquad \sum_{n=0}^{\infty} \gamma_n < +\infty$$

où l'état initial X_0 est un nombre réel arbitrairement choisi et $\theta > 0$.

1) Soit $x^* = \alpha/\theta$ et $\beta_n = (1+\theta\gamma_0)(1+\theta\gamma_1)\cdots(1+\theta\gamma_n)$. Montrer la relation de récurrence $X_{n+1} - x^* = \beta_n(X_0 - x^*) + \beta_n M_{n+1}$ avec

$$M_{n+1} = \sum_{k=0}^{n} \frac{\gamma_k}{\beta_k}\varepsilon_{k+1}.$$

2) Etudier le comportement asymptotique presque sûr de (M_n).

3) Montrer que $X_n \longrightarrow x^* + L$ p.s. avec L non nulle p.s. et conclure.

3 Records.

On peut se demander si les performances sportives seront toujours battues et si oui, à quel rythme. Afin de modéliser cette situation, on se donne une suite (X_n) de variables aléatoires indépendantes et de même loi, de densité de probabilité f. Pour tout $n \geq 1$, on note R_n le rang relatif de X_n, défini par

$$R_n = 1 + \sum_{k=1}^{n} \mathbb{I}_{(X_k > X_n)}.$$

C'est une variable aléatoire à valeurs dans $\{1, 2, \ldots, n\}$.

1) Montrer que $R_1, R_2, \ldots, R_n$ sont indépendantes avec, pour tout $n \geq 1$, $\mathbb{P}(R_n = r_n) = 1/n$ où $r_n \in \{1, 2, \ldots, n\}$.

2) Pour $n \geq 1$, si $R_n = 1$, on dit qu'il se produit un record à l'instant n. On s'intéresse à la variable aléatoire Z_n comptant le nombre de records jusqu'à l'instant n,

$$Z_n = \sum_{k=1}^{n} \mathbb{I}_{(R_k = 1)}.$$

Montrer que $Z_n / \log n \longrightarrow 1$ p.s.

3) Montrer également que

$$\frac{Z_n - \log(n)}{\sqrt{\log n}} \xrightarrow{\mathcal{L}} \mathcal{N}(0, 1).$$

4 Noyau.

Soit (X_n) une suite de variables aléatoires indépendantes et de même loi, de densité de probabilité f. On suppose que $f \in \mathcal{C}^1(\mathbb{R})$ et f' est bornée. Soit K une fonction positive, bornée, appelée noyau, telle que

$$\int_{\mathbb{R}} K(x)\,dx = 1 \qquad \text{et} \qquad \int_{\mathbb{R}} K^2(x)\,dx = \sigma^2.$$

On peut par exemple choisir le noyau uniforme $K(x) = \mathbb{I}_{(|x| \leq a)}/2a$ avec $a > 0$, le noyau d'Epanechnikov $K(x) = 3(1 - x^2/b^2)\mathbb{I}_{(|x| \leq b)}/4b$ avec $b > 0$, ou encore le noyau gaussien $K(x) = \exp(-x^2/2)/\sqrt{2\pi}$. On estime f par l'estimateur à noyau $\widehat{f}_n$ défini $\forall x \in \mathbb{R}$ par

$$\widehat{f}_n(x) = \frac{1}{n}\sum_{i=1}^{n} \frac{1}{h_i} K\Big(\frac{X_i - x}{h_i}\Big)$$

où la fenêtre $h_n = n^{-\alpha}$ avec $0 < \alpha < 1$.

1) Montrer que $\forall x \in \mathbb{R}$ fixé, $\widehat{f}_n(x) \longrightarrow f(x)$ p.s.

2) Si $1/3 < \alpha < 1$, montrer le théorème limite centrale ponctuel

$$\sqrt{nh_n}(\widehat{f}_n(x) - f(x)) \xrightarrow{\mathcal{L}} \mathcal{N}\Big(0, \frac{\sigma^2 f(x)}{1+\alpha}\Big).$$