UTILISATION DES MARTINGALES

1 Processus autorégressif.

On considère le processus autorégressif

$$X_{n+1} = \theta X_n + \varepsilon_{n+1}$$

où le paramètre inconnu $|\theta| < 1$ et l'état initial $X_0 = 0$. On suppose que (ε_n) est une suite de variables aléatoires indépendantes et de même loi $\mathcal{N}(0, \sigma^2)$ avec $\sigma^2 > 0$. On estime le paramètre inconnu θ par l'estimateur des moindres carrés

$$\widehat{\theta}_n = \frac{\sum_{k=1}^{n} X_k X_{k-1}}{\sum_{k=0}^{n-1} X_k^2}.$$

1) Montrer que l'on a l'égalité

$$\widehat{\theta}_n - \theta = \sigma^2 \frac{M_n}{<M>_n}$$

$$M_n = \sum_{k=1}^{n} X_{k-1}\varepsilon_k \qquad \text{et} \qquad <M>_n = \sigma^2 \sum_{k=1}^{n} X_{k-1}^2.$$

2) En déduire que $\widehat{\theta}_n \to \theta$ presque sûrement et que l'on a le théorème limite centrale

$$\sqrt{n}(\widehat{\theta}_n - \theta) \xrightarrow{\mathcal{L}} \mathcal{N}(0, 1 - \theta^2).$$

2 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 $n \geqslant 1$, on note R_n le rang relatif de X_n, défini par

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

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

2) Pour $n \geqslant 1$, 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,

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

Montrer que presque sûrement

$$\lim_{n\to\infty} \frac{Z_n}{\log n} = 1$$

3) Montrer que l'on a également le théorème limite centrale

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

3 Autorégressif à coefficient aléatoire.

On considère le processus autorégressif à coefficient aléatoire

$$X_{n+1} = \theta_{n+1} X_n + \varepsilon_{n+1}$$

où l'état initial $X_0 = 0$ et (θ_n) est une suite de variables aléatoires indépendantes et de même loi avec $\mathbb{E}[\theta_n] = \theta$ et $\mathbb{E}[\theta_n^2] = \tau^2 > \theta^2$. On se place dans le cas explosif avec $\tau^2 > 1$. On suppose que (ε_n) est une suite de variables aléatoires indépendantes et de même loi avec $\mathbb{E}[\varepsilon_n] = 0$ et $\mathbb{E}[\varepsilon_n^2] = \sigma^2 > 0$. On suppose également que les suites (θ_n) et (ε_n) sont indépendantes.

1) Montrer que pour tout $n \geqslant 1$, $\mathbb{E}[\theta_{n+1}\varepsilon_{n+1}|\mathcal{F}_n] = 0$,

$$\mathbb{E}[X_{n+1}|\mathcal{F}_n] = \theta X_n \qquad \text{et} \qquad \mathbb{E}[X_{n+1}^2|\mathcal{F}_n] = \tau^2 X_n^2 + \sigma^2.$$

2) Soit (Y_n) la suite donnée, pour tout $n \geqslant 1$, par $Y_n = X_n^2/\tau^{2n}$. Vérifier que (Y_n) est une sous-martingale bornée dans $\mathbb{L}^1$. En déduire que (Y_n) converge presque surement vers une variable aléatoire intégrable Y. Dans toute la suite, on supposera que Y est non nulle presque surement.

3) Montrer via le lemme de Toeplitz que

$$\lim_{n\to\infty} \frac{1}{\tau^{2n}} \sum_{k=1}^{n} X_k^2 = \left(\frac{\tau^2}{\tau^2 - 1}\right) Y \qquad \text{p.s.}$$

4) On propose d'estimer la moyenne inconnue θ par l'estimateur des moindres carrés pondérés défini par $\widetilde{\theta}_n = A_n/B_n$ avec $a_n = (1 + X_n^2)^{-1}$

$$A_n = \sum_{k=1}^{n} a_{k-1} X_k X_{k-1} \qquad \text{et} \qquad B_n = \sum_{k=1}^{n} a_{k-1} X_{k-1}^2.$$

Si $M_n = \sum_{k=1}^{n} a_{k-1} X_{k-1}(\varepsilon_k + X_{k-1}(\theta_k - \theta))$, vérifier que

$$\widetilde{\theta}_n - \theta = \frac{M_n}{B_n}.$$

5) Montrer que (M_n) est une martingale de carré intégrable dont le processus croissant $< M >_n = O(n)$ presque surement. Vérifier que $B_n/n \to 1$ puis en déduire que $\widetilde{\theta}_n \to \theta$ presque surement.

6) Si $\mathbb{E}[\theta_n^4] = \nu^4$ et $\mathbb{E}[\varepsilon_n^4] = \lambda^4$, établir le théorème limite centrale

$$\sqrt{n}(\widetilde{\theta}_n - \theta) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \tau^2 - \theta^2).$$