TP : Martingales 2 - Le Bandit à Deux Bras - Corrigé

Le présent document a été fortement inspiré du corrigé “Machines à sous” annexé au livre “Statistique en action” de Rivoirard et Stoltz. En complément on recommande la lecture de cette annexe, qui est disponible sur Internet.

1. La meilleure stratégie est de jouer toujours avec le levier A si $\theta^A > \theta^B$ et toujours avec le levier B sinon. On a alors $G_n \to \max(\theta^A, \theta^B)$ p.s. par la loi des grands nombres.

2. Soit $\mathcal{F}_n$ la filtration $\mathcal{F}_n = \sigma(X_1, \dots, X_n)$. Par définition de la notion de stratégie (non randomisée), la variable aléatoire L_{n+1} est $\mathcal{F}_n$-mesurable. Donc M_n est bien $\mathcal{F}_n$ mesurable, et elle est de carré intégrable car bornée :

$$|M_n| \leq \sum_{k=1}^{n} |X_k| + N_n^A + N_n^B \leq 2n \, .$$

De plus,

$$\mathbb{E}[M_{n+1}|\mathcal{F}_n] = M_n - \theta^A \mathbf{1}_{L_{n+1}=A} - \theta^B \mathbf{1}_{L_{n+1}=B} + \mathbb{E}[X_{n+1}|\mathcal{F}_n] \, ,$$

car $X_1, \cdots, X_n, L_1, \cdots, L_{n+1}$ sont mesurables par rapport à $\mathcal{F}_n$. Sachant $L_{n+1} = L$, X_{n+1} est une variable de Bernoulli de paramètre θ^L. Par conséquent,

$$\mathbb{E}\left[X_{n+1}|\mathcal{F}_n\right] = \theta^A \mathbf{1}_{L_{n+1}=A} + \theta^B \mathbf{1}_{L_{n+1}=B} \, .$$

On a donc bien $\mathbb{E}[M_{n+1}|\mathcal{F}_n] = M_n$ donc M_n est une martingale de carré intégrable.

Calculons le processus croissant de M_n. On a

$$(M_{n+1} - M_n)^2 = \left(X_{n+1} - \theta^A \mathbf{1}_{L_{n+1}=A} - \theta^B \mathbf{1}_{L_{n+1}=B}\right)^2 .$$

Par suite,

$$\begin{aligned}
\mathbb{E}[(M_{n+1} - M_n)^2|\mathcal{F}_n] &= \mathbb{E}\left[(X_{n+1} - \theta^A \mathbf{1}_{L_{n+1}=A} - \theta^B \mathbf{1}_{L_{n+1}=B})^2\right] \\
&= \mathbb{E}\left[(X_{n+1} - \theta^A)^2|\mathcal{F}_n\right]\mathbf{1}_{L_{n+1}=A} + \mathbb{E}\left[(X_{n+1} - \theta^B)^2|\mathcal{F}_n\right]\mathbf{1}_{L_{n+1}=B} \\
&= \theta^A(1-\theta^A)\mathbf{1}_{L_{n+1}=A} + \theta^B(1-\theta^B)\mathbf{1}_{L_{n+1}=B} \\
&\left(= \theta^{L_{n+1}}(1-\theta^{L_{n+1}})\right) .
\end{aligned}$$

Noter que ce terme est exactement la variance conditionnelle de l'accroissement $M_{n+1} - M_n$ (ici la loi conditionnelle de $M_{n+1} - M_n$ conditionnellement à $\mathcal{F}_n$ est bien une Bernoulli de paramètre $\theta^{L_{n+1}}$).

Donc en sommant

$$\langle M \rangle_n = \sum_{k=1}^{n} \mathbb{E}\left[(M_k - M_{k-1})^2|\mathcal{F}_{k-1}\right] = \theta^A(1-\theta^A)N_n^A + \theta^B(1-\theta^B)N_n^B \, .$$

3. On a

$$\min(\theta^A(1-\theta^B), \theta^B(1-\theta^B)) \cdot n \;\leq\; \langle M \rangle_n \;\leq\; N_n^A + N_n^B = n \, ,$$

d'où $\langle M \rangle_n \to +\infty$ p.s. et donc $\langle M \rangle_\infty = \infty$.

D'après la loi des grands nombres pour les martingales, on a $\dfrac{M_n}{\langle M \rangle_n} \to 0 \quad$ p.s. .

Ainsi,

$$\frac{M_n}{n} = \frac{M_n}{\langle M \rangle_n} \frac{\langle M \rangle_n}{n} \to 0 \text{ p.s. .}$$

Donc $M_n = o(n)$ presque sûrement.

On a

$$G_n = \frac{M_n}{n} + \theta^A \cdot \frac{N_n^A}{n} + \theta^B \cdot \frac{N_n^B}{n} \ ,$$

Comme $N_n^A + N_n^B = n$, on en déduit

$$\frac{M_n}{n} + \min(\theta^A, \theta^B) \leq G_n \leq \frac{M_n}{n} + \max(\theta^A, \theta^B) \ .$$

Comme $\frac{M_n}{n} \longrightarrow 0$, il s'ensuit

$$\min(\theta^A, \theta^B) \leq \liminf G_n \leq \limsup G_n \leq \max(\theta^A, \theta^B) \ .$$

4. Voir corrigé du code.

5. Voir corrigé du code.

6. Supposons $\theta^A > \theta^B$. Les quatre situations possibles à l'issue de l'étape 1 sont

(a) $L_1 = A$ et $X_1 = 1$, avec probabilité $\frac{1}{2}\theta^A \longrightarrow$ Réussite.

(b) $L_1 = A$ et $X_1 = 0$, avec probabilité $\frac{1}{2}(1 - \theta^A) \longrightarrow$ Les estimateurs sont inchangés.

(c) $L_1 = B$ et $X_1 = 1$, avec probabilité $\frac{1}{2}\theta^B \longrightarrow$ Échec.

(d) $L_1 = B$ et $X_1 = 0$, avec probabilité $\frac{1}{2}(1 - \theta^B) \longrightarrow$ Les estimateurs sont inchangés.

Dans le cas (c), pour tout $n \geq 1$, $L_n = B$ car $\widehat{\theta_n^B} > 0 = \widehat{\theta_n^A}$.

Tant que $X_i = 0$ (quel que soit L_i), on a $\widehat{\theta_n^A} = \widehat{\theta_n^B} = 0$.

Dès que $X_p = 1$, on a $\widehat{\theta_p^A} > 0$ ou $\widehat{\theta_p^B} > 0$ selon le résultat du p-ième tirage. Puis pour tout $n \geq p$, $\widehat{\theta_n^A} > 0$ alors que $\widehat{\theta_n^B} = 0$ (ou réciproquement).

La probabilité d'échec s'exprime donc par

$$\begin{aligned}
\mathbb{P}(\text{échec}) &= \mathbb{P}\Big(\exists p \in \mathbb{N}^*, X_1 = \cdots = X_{p-1} = 0, L_p = B \text{ et } X_p = 1\Big) \\
&= \sum_{p \geq 1} \mathbb{P}(X_1 = \cdots = X_{p-1} = 0, L_p = B \text{ et } X_p = 1) \\
&= \sum_{p \geq 1} \mathbb{P}(L_p = B \text{ et } X_p = 1 | X_1 = \cdots = X_{p-1} = 0)\mathbb{P}(X_1 = \cdots = X_{p-1} = 0) \\
&= \frac{\theta^B}{2} \sum_{p \geq 1} \mathbb{P}(X_1 = \cdots = X_{p-1} = 0) = \frac{\theta^B}{2} \sum_{p \geq 1} \left(\frac{1 - \theta^A}{2} + \frac{1 - \theta^B}{2} \right)^{p-1} = \frac{\theta^B}{\theta^A + \theta^B} \ .
\end{aligned}$$

Le problème vient du fait qu'on ne joue pas avec chaque levier une infinité de fois.

7. Posons

$$M_n^L = \sum_{k=1}^{n} X_k \mathbf{1}_{\{L_k = L\}} - \theta^L N_n^L \ .$$

On a bien $\widehat{\theta_n^L} - \theta^L = \frac{M_n^L}{N_n^L}$. M_n^L est $\mathcal{F}_n$-mesurable et de carré intégrable car

$$|M_n^L| \leq \sum_{k=1}^{n} |X_k| + \theta^L N_n^L \leq n + 1 \times N_n^L \leq 2n \ .$$

De plus,

$$\begin{aligned}
\mathbb{E}[M_{n+1}^L | \mathcal{F}_n] &= M_n^L + \mathbf{1}_{L_{n+1} = L} \mathbb{E}\left[X_{n+1} | \mathcal{F}_n\right] - \theta^L \mathbf{1}_{L_{n+1} = L} \\
&= M_n^L + \theta^L \mathbf{1}_{L_{n+1} = L} - \theta^L \mathbf{1}_{L_{n+1} = L} = M_n^L \ ,
\end{aligned}$$

car $X_1, \ldots, X_n, L_1, \ldots, L_{n+1}$ sont $\mathcal{F}_n$-mesurables. M_n^L est donc une martingale.

Et on peut calculer son processus croissant car

$$\begin{aligned}\mathbb{E}\left[(M_{n+1}^L - M_n^L)^2 | \mathcal{F}_n\right] &= \mathbb{E}\left[(X_{n+1}\mathbf{1}_{L_{n+1}=L} - \theta^L \mathbf{1}_{L_{n+1}=L})^2 | \mathcal{F}_n\right] \\ &= \mathbb{E}\left[(X_{n+1} - \theta^L)^2 | \mathcal{F}_n\right] \mathbf{1}_{L_{n+1}=L} \\ &= \theta^L(1-\theta^L)\mathbf{1}_{L_{n+1}=L} \ .\end{aligned}$$

Donc en sommant

$$\langle M^L \rangle_n = \theta^L(1-\theta^L) N_n^L.$$

8. Supposons $N_n^L \to +\infty$ p.s. Alors on a

$$\langle M^L \rangle_n = \theta^L(1-\theta^L) N_n^L \to +\infty \quad \text{p.s.} \ ,$$

et donc $\langle M \rangle_\infty = \infty$ p.s. D'après la loi des grands nombres pour les martingales,

$$\frac{M_n^L}{\langle M^L \rangle_n} \to 0 \quad \text{ p.s.} \ .$$

Or

$$\frac{M_n^L}{\langle M^L \rangle_n} = \frac{M_n^L}{\theta^L(1-\theta^L) N_n^L} = \frac{\widehat{\theta_n^L} - \theta^L}{\theta^L(1-\theta^L)}.$$

On a donc $\widehat{\theta_n^L} \to \theta^L$ p.s.

9. À partir de maintenant, on supposera par exemple $\theta^A > \theta^B$. Par la question précédente, on a presque sûrement

$$\widehat{\theta_n^A} \to \theta^A \quad \text{et} \quad \widehat{\theta_n^B} \to \theta^B \ .$$

En particulier, p.s. ω, il existe un rang $n_0(\omega)$ tel que

$$n \geq n_0(\omega) \quad \Longrightarrow \quad \widehat{\theta_n^A} \geq \widehat{\theta_n^B} \quad ,$$

et donc pour $n \geq n_0(\omega)$ et $n \notin \Gamma$, on a $L_{n+1} = A$.

Écrivons alors

$$G_n - \theta^A = \frac{N_n^A}{n}(\widehat{\theta_n^A} - \theta^A) + \frac{N_n^B}{n}(\widehat{\theta_n^B} - \theta^A) \ .$$

Pour $n \geq n_0(\omega)$, on a

$$n \geq N_n^A \geq n - n_0(\omega) - \mathrm{Card}(\Gamma^B \cap \{n_0(\omega), \ldots, n\}) \ ,$$

et

$$N_n^B \leq \mathrm{Card}(\Gamma^B \cap \{1, \ldots, n\}) + n_0(\omega) \ . \tag{1}$$

Puisque $\mathrm{Card}(\Gamma^L \cap \{1, \ldots, n\}) = o(n)$, on en déduit que presque sûrement,

$$\frac{N_n^A}{n} \longrightarrow 1 \quad \text{et} \quad \frac{N_n^B}{n} \longrightarrow 0 \quad \text{donc} \quad G_n \longrightarrow \theta^A = \max(\theta^A, \theta^B) \quad \text{p.s.} \ .$$

10. Voir la correction du code.

11. On va supposer maintenant que

$$\mathrm{Card}(\Gamma \cap \{1, \ldots, n\}) = o(\sqrt{n}).$$

En considérant uniquement les excitations données par le bras A, le théorème central-limite usuel donne

$$\frac{1}{\sqrt{n}} \sum_{k=1}^{n} (E_k^A - \theta^A) \xrightarrow{\text{loi}} \mathcal{N}\big(0, \theta^A(1-\theta^A)\big).$$

Or

$$\begin{aligned}
\sqrt{n}(G_n - \theta^A) &= \frac{1}{\sqrt{n}} \sum_{k=1}^{n} (X_k - \theta^A) \\
&= \frac{1}{\sqrt{n}} \sum_{k=1}^{n} \left(E_k^A - (E_k^A - E_k^B)\mathbf{1}_{L_k = B} - \theta^A \right) \\
&= \frac{1}{\sqrt{n}} \sum_{k=1}^{n} (E_k^A - \theta^A) - \frac{1}{\sqrt{n}} \sum_{k=1}^{n} (E_k^A - E_k^B)\mathbf{1}_{L_k = B}.
\end{aligned}$$

Remarquons qu'en valeur absolue, le deuxième terme est $\leq \frac{1}{\sqrt{n}}\mathrm{Card}(\Gamma^B \cap \{1, \ldots, n\})$ et donc tend vers zéro p.s. grâce aux choix de Γ^L. Le lemme de Slutsky permet de conclure que

$$\sqrt{n}(G_n - \theta^A) \xrightarrow{\text{loi}} \mathcal{N}\big(0, \theta^A(1 - \theta^A)\big).$$

12. Voir la correction du code.