Corrigé TD5. Arrêt optimal

Exercice 1. (Le probleme de moser) Il s'agit du problème d'arrêt optimal suivant. Soient $X_1, ..., X_N$ des v.a. iid positives avec fonction de répartition F et moyenne $\mathbb{E}[X_i]$ finie. On imagine connaître la loi F. Soit $\mathcal{F}_n = \sigma(X_1, ..., X_n)$ et $Y_n = X_n$: on observe en séquence N réalisations indépendantes de F, notre gain est la dernière valeur observée avant de s'arrêter. L'horizon est N: si nous ne nous arrêtons pas avant N on est obligé d'accepter le gain $Y_N = X_N$.

a) Montrer que la fonction valeur Z_n est mesurable par rapport à $\sigma(X_n)$ pour tout $1 \leqslant n \leqslant N$ (sugg: utiliser que $Y_n \hat{\in} \sigma(X_n)$ et une recurrence retrograde)

b) Soit $V_n = \mathbb{E}[Z_n]$. Montrer que $V_n = \varphi(V_{n+1})$ où $\varphi(x) = \mathbb{E}[\sup(X_1, x)]$.

c) Montrer que φ est une fonction positive et croissante, telle que $\varphi(x) - x$ est decroissante et $\varphi(x) - x \to 0$ si $x \to +\infty$.

d) Montrer que une regle optimale est

$$T^\star = \inf\{k \leqslant N : X_k \geqslant V_{N-k}\}$$

e) Soit $X_i \sim \mathcal{U}([0,1])$ pour $1 \leqslant i \leqslant N$ et $N = 6$. Montrer que la strategie optimale est donné par la procedure suivante: s'arrêter au temps 1 si $X_1 \geqslant 0.775$, s'arrêter au temps 2 si $X_2 \geqslant 0.742$, s'arrêter au temps 3 si $X_3 \geqslant 0.695$, s'arrêter au temps 4 si $X_2 \geqslant 0.625$, s'arrêter au temps 5 si $X_5 \geqslant 1/2$ ou s'arrêter à 6.

Solution.

Pour déterminer la règle d'arrêt optimal on observe qu'à cause du fait que $Y_n \in \sigma(X_n)$ on a que la fonction valeur Z_n est aussi $\sigma(X_n)$ mesurable. Démonstration par récurrence (rétrograde): c'est vrai pour $n = N$, en effet $Z_N = Y_N \in \sigma(X_N)$. Supposons qu'il est vrai pour tout $k \geqslant n+1$ et démontrons qu'il est aussi vrai pour $k = n$. De la définition de Z_n on a que $Z_n = \sup(Y_n, \mathbb{E}[Z_{n+1}|\mathcal{F}_n])$. Par indépendance des $(X_i)_i$ et par le fait qu'on suppose que $Z_{n+1} \in \sigma(X_{n+1})$ on a que $\mathbb{E}[Z_{n+1}|\mathcal{F}_n] = \mathbb{E}[Z_{n+1}]$. Donc $Z_n = \sup(Y_n, \mathbb{E}[Z_{n+1}]) \in \sigma(Y_n) = \sigma(X_n)$, ce qui permet de conclure. Donc $Z_n = \sup(X_n, \mathbb{E}[Z_{n+1}])$ pour tout $1 \leqslant n < N$ et $Z_N = X_N$. En utilisant l'identité $\max(\alpha, \beta) = \beta + (\alpha - \beta)_+$ on a que

$$\mathbb{E}[Z_n] = \mathbb{E}[\sup(Y_n, \mathbb{E}[Z_{n+1}]) = \varphi(\mathbb{E}[Z_{n+1}])$$

où $\varphi(x) = \mathbb{E}[\sup(x, X_1)] = x + \mathbb{E}[(X_1 - x)_+]$ est une fonction positive, croissante et telle que $\varphi(x) - x$ est décroissante et $\varphi(x) - x \to 0$ quand $x \to \infty$ car $\mathbb{E}[(X_1 - x)_+] \leqslant \mathbb{E}[\mathbb{I}_{X_1 > x} X_1] \to 0$ par le théorème de convergence dominée: $\mathbb{I}_{X_1 > x} X_1 \leqslant X_1$, X_1 est intégrable et $\lim_{x \to \infty} \mathbb{I}_{X_1(\omega) > x} = 0$ pour tout ω. Si l'on défini $W_1 = \mathbb{E}[X_1]$ et $W_n = \varphi(V_{n-1})$ pour $1 < n \leqslant N$ alors $\mathbb{E}[Z_N] = \mathbb{E}[X_N] = W_1$ et $\mathbb{E}[Z_n] = W_{N-n+1}$ qui est une fonction décroissante de n. Donc le gain optimal est $J_N = \mathbb{E}[Z_1] = W_N$ et la règle d'arrêt optimal est donnée par

$$T^\star = \inf\{k \leqslant N : X_k \geqslant \mathbb{E}[Z_{k+1}]\} = \inf\{k \leqslant N : X_k \geqslant W_{N-k}\}$$

Il faut donc s'arrêter dès que l'on observe une valeur de X_k supérieure au seuil W_{N-k} (qui est décroissant). La stratégie optimale demande au début d'avoir des grandes observations pour s'arrêter, mais au fur et à mesure que le temps passe le seuil d'arrêt baisse pour approcher la moyenne de X_1.

Considérons le problème de Moser avec des $X_i \sim \mathcal{U}([0,1])$. On a $\mathbb{E}[X_1]=1/2$ et

$$\mathbb{E}[(X-x)_+]=\int_0^1 (u-x)_+\mathrm{d}u=\int_x^1 (u-x)\mathrm{d}u=\frac{1-x^2}{2}-x(1-x)=\frac{1+x^2}{2}-x$$

donc $\varphi(x)=(x^2+1)/2$. Un calcul direct donne $W_1=1/2$, $W_2=0.625$, $W_3=0.695$, $W_4=0.742$, $W_5=0.775$, $W_6=0.8004$, $W_7=0.820$, $W_8=0.836$, $W_9=0.850$, $W_{10}=0.861$...

Par exemple: si $N=3$ on retrouve la stratégie optimale qu'on a déjà vu. Si $N=6$ la stratégie optimale donnée par $T^\star$ est de s'arrêter au temps 1 si $X_1 \geqslant \mathbb{E}[Z_2]=W_5=0.775$, s'arrêter au temps 2 si $X_2 \geqslant W_4=0.742$, s'arrêter au temps 3 si $X_3 \geqslant W_3=0.695$, s'arrêter au temps 4 si $X_2 \geqslant W_2=0.625$, s'arrêter au temps 5 si $X_5 \geqslant W_1=1/2$ ou s'arrêter à 6.

Exercice 2. (PROBLEME DE LA SECRETAIRE) Il s'agit de choisir parmi N objet le meilleur. On a le droit d'inspecter un objet à la fois et de décider de le choisir et donc s'arrêter ou de passer à l'inspection du suivant. Ce n'est pas possible de revenir sur ses propres pas: chaque fois on ne peut seulement garder que le dernier objet ou continuer. On veut déterminer une stratégie d'arrêt qui nous permet de maximiser la probabilité de choisir l'objet qui est le meilleur parmi les N à notre disposition. Ce problème porte le nom de « problème de la princesse » où, dans la littérature anglo-saxonne, problème classique de la secrétaire (CSP - classic secretary problem).

Le modèle mathématique est basé sur un espace d'états Ω donné par les possible permutations des N objets: $\omega \in \Omega$ est un vecteur $\omega \in \{1,\ldots,N\}^N$ tel que $\omega(i) \neq \omega(j)$ si $i \neq j$ pour tout $i, j = 1,\ldots,N$. Sur Ω on considère la distribution uniforme qui donne le même poids $1/N!$ à chaque permutation. La valeur $\omega(i)$ est le rang absolu de l'n-ème objet inspecté, donc si $\omega(i)=1$ le meilleur objet se trouve dans la position i. On remarque qu'on ne peut pas observer directement les ω (on ne connaît pas le classement des objets jusqu'a ce qu'on ai inspecté tous les N objets). A chaque pas n on observe une variable $X_n(\omega)$ qui donne le rang *relatif* de l'n-ème objet inspecté par rapport à tous les $n-1$ objets inspectés auparavant. Donc $X_1=1$, $X_2 \in \{1,2\}$,..., $X_n \in \{1,\ldots,n\}$ et $X_N(\omega)=\omega(N)$: une fois que j'ai inspecté tous les objets je connais leur classement absolu. A chaque instant n je connais $\mathcal{F}_n=\sigma(X_1,\ldots,X_n)$ la tribu engendrée par le rangs relatifs des premiers n objets. Exemple: si $N=4$ et $\omega=(3,4,1,2)$ alors $X_1(\omega)=1, X_2(\omega)=2, X_3(\omega)=1$, $X_4(\omega)=2$. Soit $\Xi=\{(x_1,\ldots,x_N)\in\mathbb{N}^N\colon 1\leqslant x_k\leqslant k, k=1,\ldots,N\}$ l'ensemble des valeurs possibles pour le vecteur aléatoire $X=(X_1,\ldots,X_N)$. On remarque que l'application $X\colon\Omega\to\Xi$ que envoie chaque possible permutation des N objet vers la correspondante suite des rangs relatifs est bijective, i.e. existe $\Psi\colon \Xi\to\Omega$ telle que $\Psi(X(\omega))=\omega$. Ce qu'il est équivalent à dire que donné la suite des rangs relatifs $x_1,\ldots,x_N$ on peut reconstruire le valeurs de $\omega_1,\omega_2,\ldots,\omega_N$.

a) Montrer que pour tout $(x_1,\ldots,x_N)\in\Xi$ on a

$$\mathbb{P}(X_1=x_1, X_2=x_2,\ldots,X_N=x_N)=1/N!$$

b) Monter que pour tout $n\leqslant N$ on a que $\mathbb{P}(X_n=j)=1/n$ pour $j=1,\ldots,n$ et que les v.a. $X_1,\ldots,X_n$ sont indépendantes.

c) L'objectif est de trouver une stratégie d'arrêt (donné par un t.a.) qui nous permet de optimiser la probabilité de choisir l'objet meilleur parmi les N disponibles. Autrement dit on veut maximiser $\mathbb{P}(\omega(T)=1)=\mathbb{E}[1_{\omega(T)=1}]$ pour tout T t.a. de la filtration $\mathcal{F}$ et borné par N. On défini un processus adapté Y par $Y_k=\mathbb{E}[1_{\omega(k)=1}|\mathcal{F}_k]\in\mathcal{F}_k$. Montrer que

$$\mathbb{P}(\omega(T)=1)=\mathbb{E}[Y_T].$$

d) Montrer que $Y_n=1_{X_n=1}\frac{n}{N}$ et donc que $Y_n\hat{\in}\sigma(X_n)$.

e) Montrer que $Z_n \hat{\in} \sigma(X_n)$ et donc que un temps d'arrêt optimal est donné par

$$T^\star = \inf\{k \leqslant N \colon \mathbb{E}(Z_{k+1}) \leqslant k/N, X_k = 1\}$$

f) Montrer que $\mathbb{E}[Z_n]$ est une fonction decroissante de n et donc que il existe r tel que

$$T^\star = T_r = \inf\{r \leqslant k \leqslant N \colon X_k = 1\} \cup \{N\}$$

g) Montrer que pour tout $1 \leqslant r \leqslant N$:

$$G_N(r) = \mathbb{E}[Y_{T_r}] = \mathbb{P}(\omega(T_r) = 1) = \frac{r-1}{N} \sum_{k=r}^{N} \frac{1}{k-1}$$

h) Montrer que $\lim_{N \to \infty} G_N(xN) = -x \log x$ et que cette fonction à un maximum pour $x = 1/e \simeq 0.37$. Donc dans la limite d'un grand nombre d'objects la strategie optimale est de ne laisser defiler une proportion du 37% et apres choisir le premier meilleur de tout les precedents.

Solution.

Soit $\Xi = \{(x_1, \ldots, x_N) \in \mathbb{N}^N \colon 1 \leqslant x_k \leqslant k, k = 1, \ldots, N\}$ l'ensemble des valeurs possibles pour le vecteur aléatoire $X = (X_1, \ldots, X_N)$. On remarque que la taille de Ξ est $N!$. On commence par montrer que pour tout $(x_1, \ldots, x_N) \in \Xi$ on a

$$\mathbb{P}(X_1 = x_1, X_2 = x_2, \ldots, X_N = x_N) = 1/N!$$

En effet l'application $X \colon \Omega \to \Xi$ que envoie chaque possible permutation des N objet vers la correspondante suite des rangs relatifs est bijective, i.e. existe $\Psi \colon \Xi \to \Omega$ telle que $\Psi(X(\omega)) = \omega$. Ce qu'il est équivalent à dire que donné la suite des rangs relatifs $x_1, \ldots, x_N$ on peut reconstruire le valeurs de $\omega_1, \omega_2, \ldots, \omega_N$. Donc $\mathbb{P}(X = x) = \mathbb{P}(\Psi(X) = \Psi(x)) = \mathbb{P}(\omega = \Psi(x)) = 1/N!$ car ω est uniforme sur Ω et $\mathrm{Card}(\Omega) = N!$. Maintenant on a aussi

$$\mathbb{P}(X_n = j) = \sum_{x_1=1}^{1} \sum_{x_2=1}^{2} \cdots \sum_{x_{n-1}=1}^{n-1} \sum_{x_{n+1}=1}^{n+1} \ldots \sum_{x_N=1}^{N} \mathbb{P}(X_1 = x_1, X_2 = x_2, \ldots, X_N = x_N) = \frac{1}{n}$$

et donc $\mathbb{P}(X_1 = x_1, X_2 = x_2, \ldots, X_N = x_N) = \mathbb{P}(X_1 = x_1) \cdots \mathbb{P}(X_N = x_N)$ qui montre l'indépendance.

On doit d'abord récrire notre critère à optimiser dans la forme d'une espérance d'une v.a. $\mathcal{F}_T$ mesurable:

$$\mathbb{E}[1_{\omega(T)=1}] = \sum_{k=1}^{N} \mathbb{E}[1_{\omega(T)=1} \cdot 1_{T=k}] = \sum_{k=1}^{N} \mathbb{E}[1_{\omega(k)=1} \cdot 1_{T=k}] = \sum_{k=1}^{N} \mathbb{E}[\mathbb{E}[1_{\omega(k)=1} | \mathcal{F}_k] \cdot 1_{T=k}]$$

où on a utilisé le fait que $\{T = k\} \in \mathcal{F}_k$ par définition de t.a. et les propriétés de l'espérance conditionnelle. Si l'on défini un processus adapté Y par $Y_k = \mathbb{E}[1_{\omega(k)=1} | \mathcal{F}_k] \in \mathcal{F}_k$, l'on a

$$\mathbb{E}[1_{\omega(T)=1}] = \sum_{k=1}^{N} \mathbb{E}[Y_k \cdot 1_{T=k}] = \sum_{k=1}^{N} \mathbb{E}[Y_T \cdot 1_{T=k}] = \mathbb{E}[Y_T]$$

et $Y_T \in \mathcal{F}_T$: le critère adapté qui nous cherchions. On est donc dans le cadre du théorème précédent: horizon fini N, fonction de gain $Y_n \in [0, 1]$ et donc intégrable pour tout $0 \leqslant n \leqslant N$. La solution du problème d'optimisation est donné par $T^\star = \inf\{k \leqslant N \colon Y_k = Z_k\}$ où Z est l'enveloppe de Snell de Y. Il nous reste donc à calculer cette fonction valeur et expliciter le temps d'arrêt $T^\star$ comme fonction de $X_1, \ldots, X_{T^\star}$ (qui sont les quantités qu'on observe pratiquement).

On commence par expliciter le critère $Y_n = \mathbb{P}(\omega(n) = 1|\mathcal{F}_n)$. On remarque que l'événement $\{\omega(n) = 1\}$ est équivalent à $\{X_n = 1, X_{n+1} \neq 1, \ldots, X_N \neq 1\}$ et par l'indépendance des $(X_k)_{1 \leqslant k \leqslant N}$:

$$Y_n = \mathbb{P}(X_n = 1, X_{n+1} \neq 1, \ldots, X_N \neq 1|\mathcal{F}_n) = \mathbb{P}(X_n = 1|\mathcal{F}_n)\mathbb{P}(X_{n+1} \neq 1)\cdots\mathbb{P}(X_N \neq 1)$$

$$= 1_{X_n=1} \cdot \frac{n}{n+1} \cdot \ldots \cdot \frac{N-1}{N} = 1_{X_n=1}\frac{n}{N}.$$

donc $Y_n \in \sigma(X_n)$. Cette propriété entraîne que Z_{n+1} est indépendante de $\mathcal{F}_n$: en effet $Z_N = Y_N = 1_{X_N=1} \in \sigma(X_N)$, $Z_{N-1} = \sup(Y_{N-1}, \mathbb{E}(Z_N|\mathcal{F}_{N-1})) \in \sigma(X_{N-1})$ et par induction on obtient $Z_n \in \sigma(X_n)$ donc, par indépendance des X on a bien que $Z_{n+1} \in \sigma(X_{n+1}) \perp \mathcal{F}_n$. Cela implique qu'on a $Z_n = \sup(Y_n, \mathbb{E}(Z_{n+1}))$ et qui

$$T^\star = \inf\{k \leqslant N: \mathbb{E}(Z_{k+1}) \leqslant Y_k\} = \inf\{k \leqslant N: \mathbb{E}(Z_{k+1}) \leqslant k/N, X_k = 1\}$$

en effet on a toujours $\mathbb{E}(Z_n) \geqslant \mathbb{E}(Z_N) = \mathbb{P}(X_N = 1) = 1/N > 0$ par la propriété de sur-martingale de Z. L'espérance de Z_n est donc décroissante en n et la stratégie est d'attendre que $\mathbb{E}(Z_{k+1})$ tombe au dessous du seuil k/N et après de s'arrêter sur le premier objet meilleur de tous les autres vus auparavant. En effet si on appelle r le premier entier $\leqslant N$ tel que $\mathbb{E}(Z_{r+1}) \leqslant r/N$, on a que $\mathbb{E}(Z_k) \geqslant \mathbb{E}(Z_r) > (r-1)/N > k/N$ pour tout $k < r$ et que $\mathbb{E}(Z_k) \leqslant \mathbb{E}(Z_{r+1}) \leqslant r/N \leqslant k/N$ pour tout $k \geqslant r$:

$$T^\star = \inf\{k \in [r, N]: X_k = 1\} =: T_r$$

la stratégie optimale est donnée par le temps d'arrêt T_r: attendre le premier instant k après r où on observe $X_k = 1$.

Ayant établi que la règle optimale est de la forme T_r il nous reste à déterminer $r \in [1, N]$ de façon telle que $\mathbb{E}[Y_{T_r}]$ soit maximale. Cela est équivalent a maximiser $\mathbb{P}(\omega(T_r) = 1)$ (car en effet on a déjà montre que pour tout t.a. $\mathbb{P}(\omega(T) = 1) = \mathbb{E}[Y_T]$). Or

$$\mathbb{P}(\omega(T_r) = 1) = \sum_{k=1}^{N} \mathbb{P}(\omega(T_r) = 1, T_r = k) = \sum_{k=1}^{N} \mathbb{P}(\omega(k) = 1, T_r = k)$$

$$= \sum_{k=r}^{N} \mathbb{P}(X_r \neq 1, \ldots, X_{k-1} \neq 1, X_k = 1, X_{k+1} \neq 1, \ldots, X_N \neq 1)$$

$$= \sum_{k=r}^{N} \mathbb{P}(X_r \neq 1)\cdots\mathbb{P}(X_{k-1} \neq 1)\mathbb{P}(X_k = 1)\mathbb{P}(X_{k+1} \neq 1)\cdots\mathbb{P}(X_N \neq 1)$$

où on a utilise le fait que $\{\omega(k) = 1, T_r = k\} = \{X_r \neq 1, \ldots, X_{k-1} \neq 1, X_k = 1, X_{k+1} \neq 1, \ldots, X_N \neq 1\}$ et l'indépendance des X_k. Donc

$$G_r = \mathbb{P}(\omega(T_r) = 1) = \sum_{k=r}^{N} \frac{r-1}{r} \cdots \frac{k-2}{k-1} \frac{1}{k} \frac{k}{k+1} \cdots \frac{N-1}{N} = \frac{r-1}{N} \sum_{k=r}^{N} \frac{1}{k-1}$$

est le gain moyen de la stratégie T_r pour tout $r \in [1, N]$. La stratégie optimale est donc donné par le $r \in [1, N]$ qui maximise la fonction G_r. Si l'on note r_* la valeur optimale on peut facilement calculer la table suivante

n	1	2	3	4	5	6	7	8
r_*	1	1	2	2	3	3	3	4
G_{r_*}	1.0	0.5	0.5	0.458	0.433	0.428	0.414	0.41

Dans la limite où $N \to \infty$ et $r/N = x \in (0,1)$ on a

$$G_r = \frac{Nx-1}{N} \sum_{k=Nx}^{N} \frac{1}{k-1} \simeq x \int_x^1 \frac{\mathrm{d}x}{x} = -x \log x$$

cette fonction a un maximum pour $\log x = -1$ et donc pour $x = 1/e \simeq 0.368$. La valeur asymptotique de G_{Nx} est aussi $\simeq 0.368$. La stratégie optimale est donc de laisser défiler les premiers $r \simeq N/e$ objets et après de choisir le premier qu'on trouve le meilleur. De cette façon on a une probabilité de $\simeq 36.8\%$ de tomber sur l'objet de rang maximal.