

Corrigé TD5. Arrêt optimal

Exercice 1. (Le problème de Moser) Il s'agit du problème d'arrêt optimal suivant. Soient X_1, \dots, X_N des v.a. iid positives avec fonction de répartition F et moyenne $E[X_i]$ finie. On imagine connaître la loi F . Soit $F_n = (X_1, \dots, X_n)$ 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. Le horizon est N : si nous ne nous arrêtons pas avant N on est obligé d'accepter le gain $Y_N = X_N$.

- Montrer que la fonction valeur Z_n est mesurable par rapport à $\mathcal{F}_n(X_n)$ pour tout $1 \leq n < N$ (sugg: utiliser que $Y_n \in \mathcal{F}_n(X_n)$ et une récurrence rétrograde)
- Soit $V_n = E[Z_n]$. Montrer que $V_n = (V_{n+1}) \circ \mathcal{F}_n(x) = E[\sup(X_1, x)]$.
- Montrer que V_n est une fonction positive et croissante, telle que $V_n(x)$ est décroissante et $V_n(x) \geq 0$ si $x \geq 0$.
- Montrer que une règle optimale est

$$T = \inf \{k \leq N: X_k \geq V_{N-k+1}\}$$

- Soit $X_i \sim U([0, 1])$ pour $1 \leq i \leq N$ et $N = 6$. Montrer que la stratégie optimale est donnée par la procédure suivante: s'arrêter au temps 1 si $X_1 \geq 0.775$, s'arrêter au temps 2 si $X_2 \geq 0.742$, s'arrêter au temps 3 si $X_3 \geq 0.695$, s'arrêter au temps 4 si $X_4 \geq 0.625$, s'arrêter au temps 5 si $X_5 \geq 1/2$ ou s'arrêter à 6.

Solution.

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

$$E[Z_n] = E[\sup(X_n, E[Z_{n+1}])] = E[X_n] + E[(E[Z_{n+1}] - X_n)_+]$$

où $\phi(x) = E[\sup(x, X_1)] = x + E[(X_1 - x)_+]$ est une fonction positive, croissante et telle que $\phi(x)$ est décroissante et $\phi(x) \geq 0$ quand $x \geq 0$ car $E[(X_1 - x)_+] = E[X_1 - x | X_1 > x] P(X_1 > x) \geq 0$ par le théorème de convergence dominée. $I_{X_1 > x} X_1$, X_1 est intégrable et $\lim_{x \rightarrow \infty} I_{X_1 > x} X_1 = 0$ pour tout x . Si on définit $W_1 = E[X_1]$ et $W_n = (V_{N-k+1})$ pour $1 < n \leq N$ alors $E[Z_N] = E[X_N] = W_1$ et $E[Z_n] = W_{N-k+1}$ qui est une fonction décroissante de n . Donc le gain optimal est $J_N = E[Z_1] = W_N$ et la règle d'arrêt optimale est donnée par

$$T = \inf \{k \leq N: X_k \geq E[Z_{N-k+1}]\} = \inf \{k \leq N: X_k \geq W_{N-k+1}\}$$

Il faut donc s'arrêter dès que l'on observe une valeur de X_k supérieure au seuil W_{N-k+1} (qui est décroissant). La stratégie optimale demande au début d'attendre des grandes observations pour s'arrêter, mais au fur et à mesure que le temps passe le seuil diminue pour approcher la moyenne de X_1 .

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

$$E[(X \otimes x)_+] = \int_0^1 (u \otimes x)_+ du = \int_x^1 (u \otimes x) du = \frac{1 \otimes x^2}{2} \otimes x(1 \otimes x) = \frac{1 + x^2}{2} \otimes x$$

donc $\phi(x) = (x^2 + 1)/2$. 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 vue. Si $N = 6$ la stratégie optimale donnée par T est de sélectionner au temps 1 si $X_1 \leq E[Z_2] = W_5 = 0.775$, sélectionner au temps 2 si $X_2 \leq W_4 = 0.742$, sélectionner au temps 3 si $X_3 \leq W_3 = 0.695$, sélectionner au temps 4 si $X_4 \leq W_2 = 0.625$, sélectionner au temps 5 si $X_5 \leq W_1 = 1/2$ ou sélectionner 6.

Exercice 2. (problème de la secrétaire) 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 sélectionner 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 déterministe 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 dans la littérature anglo-saxonne, problème classique de la secrétaire (CSP - classic secretary problem).

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

a) Montrer que pour tout $(x_1, \dots, x_N) \in \Omega$ on a

$$P(X_1 = x_1, X_2 = x_2, \dots, X_N = x_N) = 1/N!$$

b) Montrer que pour tout $n \leq N$ on a que $P(X_n = j) = 1/n$ pour $j = 1, \dots, n$ et que les v.a. X_1, \dots, X_N sont indépendantes.

c) L'objectif est de trouver une stratégie déterministe (donnée 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 $P(\phi(T) = 1) = E[1_{\phi(T)=1}]$ pour tout T t.a. de la tribu \mathcal{F} et borné par N . On définit un processus adapté Y par $Y_k = E[1_{\phi(k)=1}/F_k] | \mathcal{F}_k$. Montrer que

$$P(\phi(T) = 1) = E[Y_T].$$

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

e) Montrer que $Z_n \leq (X_n)$ et donc que un temps d'arrêt optimal est donné par

$$T = \inf \{k \leq N : E(Z_{k+1}) \leq k/N, X_k = 1\}$$

f) Montrer que $E[Z_n]$ est une fonction décroissante de n et donc que il existe r tel que

$$T = T_r = \inf \{r \leq k \leq N : X_k = 1\} \leq \{N\}$$

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

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

h) Montrer que $\lim_N G_N(xN) = x \log x$ et que cette fonction a un maximum pour $x = 1/e \approx 0.37$. Donc dans la limite d'un grand nombre d'objets la stratégie optimale est de ne laisser de côté une proportion du 37 % et après choisir le premier meilleur de tout les précédents.

Solution.

Soit $\Omega = \{(X_1, \dots, X_N) \in \mathbb{N}^N : 1 \leq X_k \leq N, k = 1, \dots, N\}$ l'ensemble des valeurs possibles pour le vecteur aléatoire $X = (X_1, \dots, X_N)$. On remarque que la taille de Ω est $N!$. On commence par montrer que pour tout $(x_1, \dots, x_N) \in \Omega$ on a

$$P(X_1 = x_1, X_2 = x_2, \dots, X_N = x_N) = 1/N!$$

En effet l'application $X \mapsto$ qui envoie chaque possible permutation des N objet vers la correspondante suite des rangs relatifs est bijective, i.e. existe σ telle que $(X(\sigma)) = (x_1, \dots, x_N)$. Ce qui est équivalent à dire que donnée la suite des rangs relatifs x_1, \dots, x_N on peut reconstruire les valeurs de x_1, x_2, \dots, x_N . Donc $P(X = x) = P(\sigma(X) = (x)) = P(\sigma = (x)) = 1/N!$ car σ est uniforme sur Ω et $\text{Card}(\Omega) = N!$. Maintenant on a aussi

$$P(X_N = j) = \frac{1}{N} \quad \dots \quad P(X_1 = x_1, X_2 = x_2, \dots, X_N = x_N) = \frac{1}{N}$$

et donc $P(X_1 = x_1, X_2 = x_2, \dots, X_N = x_N) = P(X_1 = x_1) \dots P(X_N = x_N)$ qui montre l'indépendance.

On doit donc réécrire notre critère d'optimisation dans la forme d'une espérance d'une v.a. mesurable: F_T

$$E[1_{\{T=1\}}] = \sum_{k=1}^N E[1_{\{T=1\}} 1_{\{X_{\sigma(k)}=k\}}] = \sum_{k=1}^N E[1_{\{X_{(k)}=1\}} 1_{\{X_{\sigma(k)}=k\}}] = \sum_{k=1}^N E[E[1_{\{X_{(k)}=1\}}/F_k] 1_{\{X_{\sigma(k)}=k\}}]$$

On a utilisé le fait que $\{T=k\} \subset F_k$ par définition de t.a. et les propriétés de l'espérance conditionnelle. Si l'on définit un processus adapté Y par $Y_k = E[1_{\{X_{(k)}=1\}}/F_k] \mid F_k$, l'on a

$$E[1_{\{T=1\}}] = \sum_{k=1}^N E[Y_k 1_{\{X_{\sigma(k)}=k\}}] = \sum_{k=1}^N E[Y_T 1_{\{X_{\sigma(k)}=k\}}] = E[Y_T]$$

et $Y_T = F_T$: le critère adapté qui nous cherchions. On est donc dans le cadre du théorème précédent: horizon N , fonction de gain $Y_n \in [0, 1]$ et donc intégrable pour tout $0 \leq n \leq N$. La solution du problème d'optimisation est donnée par $T = \inf \{k \leq N : Y_k = Z_k\}$ 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 comme fonction de X_1, \dots, X_T (qui sont les quantités que l'on observe pratiquement).

On commence par expliciter le critère $Y_n = P((\tau) = 1/F_n)$. On remarque que l'événement $\{(\tau) = 1\}$ est équivalent à $\{X_n = 1, X_{n+1} = 1, \dots, X_N = 1\}$ et par indépendance des $(X_k)_{1 \leq k \leq N}$:

$$\begin{aligned} Y_n &= P(X_n = 1, X_{n+1} = 1, \dots, X_N = 1/F_n) = P(X_n = 1/F_n)P(X_{n+1} = 1) \dots P(X_N = 1) \\ &= 1_{X_n=1} \frac{1}{n+1} \frac{1}{n} \dots \frac{1}{N} = 1_{X_n=1} \frac{1}{N}. \end{aligned}$$

donc $Y_n = 1/N$ (X_n). Cette propriété entraîne que Z_{n+1} est indépendante de F_n ; en effet $Z_N = Y_N = 1_{X_N=1} (X_N)$, $Z_N = \sup(Y_N, E(Z_N/F_N)) = (X_N)$ et par induction on obtient $Z_n = 1_{X_n=1} (X_n)$ donc, par indépendance des X on a bien que $Z_{n+1} = 1_{X_{n+1}=1} (X_{n+1}) \perp F_n$. Cela implique que $Z_n = \sup(Y_n, E(Z_{n+1}))$ et qui

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

en effet on a toujours $E(Z_n) = E(Z_N) = P(X_N = 1) = 1/N > 0$ 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 $E(Z_{k+1})$ tombe au dessous du seuil k/N et après de sélectionner sur le premier objet meilleur de tous les autres vus auparavant. En effet et si on appelle r le premier entier $\leq N$ tel que $E(Z_{r+1}) \leq r/N$, on a que $E(Z_k) = E(Z_r) > (r-1)/N > k/N$ pour tout $k < r$ et que $E(Z_k) = E(Z_{r+1}) \leq r/N \leq k/N$ pour tout $k \geq r$.

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

la stratégie optimale est donc 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 $E[Y_{T_r}]$ soit maximale. Cela est équivalent à maximiser $P((T_r) = 1)$ (car en effet on a démontré que pour tout t.a. $P((\tau) = 1) = E[Y_T]$). Or

$$\begin{aligned} P((T_r) = 1) &= \sum_{k=1}^N P((T_r) = 1, T_r = k) = \sum_{k=1}^N P((X_k) = 1, T_r = k) \\ &= \sum_{k=r}^N P(X_r = 1, \dots, X_{k-1} = 1, X_k = 1, X_{k+1} = 1, \dots, X_N = 1) \\ &= \sum_{k=r}^N P(X_r = 1) \dots P(X_{k-1} = 1) P(X_k = 1) P(X_{k+1} = 1) \dots P(X_N = 1) \end{aligned}$$

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

$$G_r = P((T_r) = 1) = \sum_{k=r}^N \frac{1}{r} \frac{1}{k} \frac{1}{k+1} \dots \frac{1}{N} = \frac{1}{N} \sum_{k=r}^N \frac{1}{k}$$

est le gain moyen de la stratégie T_r pour tout $r \in [1, N]$. La stratégie optimale est donc donnée 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 $N \rightarrow \infty$ et $r/N = x \in (0, 1)$ on a

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

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