Stochastic Process (ENPC)
Monday, 27th of January 2025 (2h30)

Vocabulary (english/*français*): positive=*strictement positif*; cumulative distribution function =*fonction de répartition*. $[0, 1) =$ *[0,1[*.

Exercise 1 (Renewal process)**.** We consider a machine working during a positive period of D_1 days until its breaks down, and then it is immediately replaced by a new machine which will be working for a positive period of D_2 days and so on. Set $S_0 = 0$ and $S_k = \sum_{\ell=1}^{k} D_\ell$ for $k \in \mathbb{N}^*$; so that S_k is the replacement day of the k-th machine. We consider the associated age process $Y = (Y_n, n \in \mathbb{N})$ of the current machine at work, see Fig. 1, defined by:

$$Y_n = n - \sup\{S_k : \ S_k \leq n\}.$$

Notice that when a new machine is installed, its age is equal to zero: $Y_{S_k} = 0$ for all $k \in \mathbb{N}$.

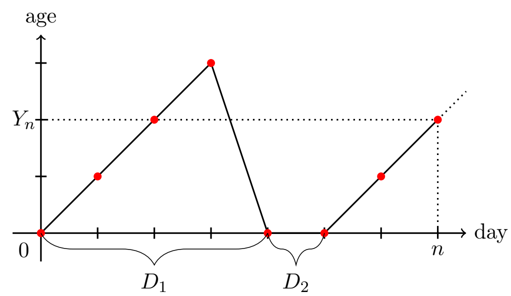

Figure 1: A realization of the age process $n \mapsto Y_n$ (with $D_1 = 4$, $D_2 = 1$ and $D_3 > 2$).

Let D be an unbounded $\mathbb{N}^*$-valued random variable. We assume that the random variables $(D_\ell, \ell \in \mathbb{N}^*)$ are independent and distributed as D. Set $\rho = (\rho(\ell) = \mathbb{P}(D > \ell \,|\, D \geq \ell), \ell \in \mathbb{N}^*)$.

1. Let $n \in \mathbb{N}$ and $\mathbf{y}_n = (y_0 = 0, y_1, \ldots, y_{n+1})$ such that $y_{\ell+1} \in \{0, y_\ell + 1\}$ for $\ell \in \{0, \ldots, n\}$.

   (a) Using $\Delta = \{\ell \in \{1, \ldots, n\} : \ y_\ell = 0\}$ and $k = \operatorname{card} \Delta$, check that:

   $$\mathbb{P}(Y_0 = y_0, \ldots, Y_n = y_n) = \mathbb{P}(\{S_1, \ldots, S_k\} = \Delta \quad \text{and} \quad D_{k+1} \geq 1 + y_n),$$

   and:

   $$\mathbb{P}(Y_0 = y_0, \ldots, Y_n = y_n, Y_{n+1} = 0) = \mathbb{P}(\{S_1, \ldots, S_k\} = \Delta \quad \text{and} \quad D_{k+1} = 1 + y_n).$$

   (b) Prove there exists a stochastic matrix P (which can be written using ρ and which does not depend on n and $\mathbf{y}_n$) such that:

   $$\mathbb{P}(Y_0 = y_0, \ldots, Y_n = y_n, Y_{n+1} = y_{n+1}) = \mathbb{P}(Y_0 = y_0, \ldots, Y_n = y_n)\, P(y_n, y_{n+1}).$$

   (c) Deduce that Y is a Markov chain.

2. (a) Check that $\prod_{\ell=1}^{n} \rho(\ell) = \mathbb{P}(D \geq n + 1)$ for all $n \in \mathbb{N}^*$.

   (b) Is the Markov chain Y irreducible?

3. Prove that Y is recurrent, and give a necessary and sufficient condition for Y to be positive recurrent.

4. When Y is positive recurrent, compute the invariant probability using the mean and the cumulative distribution of D.

$\triangle$

Exercise 2 (Searching a parking place)**.** You still take your car to go to the opera, and you want to park in the opera street as close as possible to the entrance of the opera. The opera street is an infinite a one-way street (you can only go forward!) with parking places labeled by $\mathbb{N}^*$, starting with the place 1, and the opera is in front of the n_0-th place, with $n_0 > 1$. Set $X_n = 0$ if the n-th place is empty and $X_n = 1$ otherwise.

At step n you are in front of the n-th place (unless you have already parked). If the place is empty then you can park there which corresponds to a loss (or cost) $|n_0 - n|$ given by the distance from this place to the opera; if the place is not empty then you can not park and for this reason the loss is set to be infinity. So, with the convention that $\infty * 0 = 0$, the loss at step n is:

$$L_n = |n_0 - n|(1 - X_n) + \infty * X_n.$$

We assume that the random variables $(X_n, n \in \mathbb{N}^*)$ are independent Bernoulli with known parameter $p \in (0, 1)$. Set $q = 1 - p$. For $n \in \mathbb{N}^*$, consider the σ-field $\mathcal{F}_n = \sigma(X_1, \ldots, X_n)$ corresponding to your observation of the first n places. Let $\mathcal{T}$ denote the set of stopping times with respect to the filtration $\mathbb{F} = (\mathcal{F}_n, n \in \mathbb{N}^*)$. The aim of the problem is to find the minimal loss $\mathbb{E}[L_\tau]$ over all the stopping times $\tau \in \mathcal{T}$, and determine, if it exists, an optimal stopping time τ', that is, a stopping time such that:

$$\mathbb{E}[L_{\tau'}] = \inf_{\tau \in \mathcal{T}} \mathbb{E}[L_\tau].$$

To do so, we consider the minimal loss at the n-step (assuming we didn't park before) $S_n = \operatorname{essinf}_{\tau \in \mathcal{T}_n} \mathbb{E}[L_\tau \,|\, \mathcal{F}_n]$, where $\mathcal{T}_n$ is the set of stopping times $\tau \in \mathcal{T}$ such that $\tau \geq n$.

1. What is the best strategy if $n \geq n_0$. Deduce that $S_{n_0} = X_{n_0}/q$. (Recall $q = 1 - p$.)

2. Recall the optimal equations satisfied by $(S_n, n \in \mathbb{N}^*)$. Prove that for $n \in \{1, \ldots, n_0\}$:
$$S_n = \alpha_n X_n + \min(n_0 - n, \alpha_n)(1 - X_n), \tag{1}$$
where $\alpha_{n_0} = 1/q$ and for $n \in \{1, \ldots, n_0\}$:
$$\alpha_n = p\alpha_{n+1} + q \min(n_0 - n - 1, \alpha_{n+1}). \tag{2}$$

3. Check there exists optimal stopping times and recall the expression of one of them. Check all of them are a.s. finite.

4. Assume $p \leq 1/2$.
   (a) Prove that $\alpha_n = p/q$ for all $n \in \{1, \ldots, n_0 - 1\}$.
   (b) Prove that $S_n = p/q$ for all $n \in \{1, \ldots, n_0 - 1\}$ and explicit an optimal stopping time.

5. We consider the general case $p \in (0, 1)$. Set $n_* = \inf\{n \in \{1, \ldots, n_0\} : \alpha_n \geq n_0 - n\}$.
   (a) Prove that if $\alpha_{n+1} \leq n_0 - n - 1$ then $\alpha_n < n_0 - n$, and then that $\alpha_n > n_0 - n$ for all $n \in \{n_* + 1, \ldots, n_0\}$.
   (b) Deduce that the following stopping time is optimal:
   $$\tau_* = \inf\{n \geq n_* : X_n = 0\}.$$
   (c) Prove that $\alpha_n = n_0 - n + q^{-1}(2p^{n_0 - n} - 1)$ for $n \in \{n^* - 1, \ldots, n_0\} \cap \mathbb{N}^*$.
   (d) Check that if $p \in (0, 1)$ belongs to $[2^{-1/r}, 2^{-1/(r+1)})$ for $r \in \{0, \ldots, n_0 - 2\}$, then $n_* = n_0 - r$, and if p belongs to $[2^{-1/(n_0 - 1)}, 1)$ then $n_* = 1$.

6. (*Open question.*) How to model the problem if p is unknown (and still use the optimal stopping time framework to solve it)?

$\triangle$

Correction

Exercise 1

1. (a) By construction of the process Y and the definition of $\mathbf{y}_n$, we have:

$$\{Y_0 = y_0, \dots, Y_n = y_n\} = \Big\{\{S_A, \dots, S_k\} = \Delta \quad \text{and} \quad D_{k+1} \geq 1 + y_n\Big\},$$
$$\{Y_0 = y_0, \dots, Y_n = y_n, Y_{n+1} = 0\} = \Big\{\{S_1, \dots, S_k\} = \Delta \quad \text{and} \quad D_{k+1} = 1 + y_n\Big\}.$$

(b) We have:

$$\begin{aligned}
&\mathbb{P}(Y_0 = y_0, \dots, Y_n = y_n, Y_{n+1} = 0)\\
&\qquad = \mathbb{P}(\{S_1, \dots, S_k\} = \Delta \quad \text{and} \quad D_{k+1} = 1 + y_n)\\
&\qquad = \mathbb{P}(\{S_1, \dots, S_k\} = \Delta \quad \text{and} \quad D_{k+1} \geq 1 + y_n)\\
&\qquad\qquad \mathbb{P}(D_{k+1} = 1 + y_n \,|\, \{S_1, \dots, S_k\} = \Delta \quad \text{and} \quad D_{k+1} \geq 1 + y_n)\\
&\qquad = \mathbb{P}(\{S_0, \dots, S_k\} = \Delta \quad \text{and} \quad D_{k+1} \geq 1 + y_n)\,(1 - \rho(y_n))\\
&\qquad = \mathbb{P}(Y_0 = y_0, \dots, Y_n = y_n)\,(1 - \rho(y_n)),
\end{aligned}$$

where we used that D_{k+1} is independent of $D_1, \dots, D_k$ and thus of $S_1, \dots, S_k$ for the third equality. Since $Y_{n+1} \in \{0, Y_n + 1\}$ a.s., we deduce that:

$$\mathbb{P}(Y_0 = y_0, \dots, Y_n = y_n, Y_{n+1} = y_{n+1}) = \mathbb{P}(Y_0 = y_0, \dots, Y_n = y_n)\, P(y_n, y_{n+1})$$

with $P(\ell, 0) = 1 - \rho(1+\ell)$, $P(\ell, \ell+1) = \rho(1+\ell)$ and $P(\ell, \ell') = 0$ for all $\ell' \in \mathbb{N} \setminus \{0, \ell+1\}$ and for all $\ell \in \mathbb{N}$.

(c) By backward recursion, we deduce that $\mathbb{P}(Y_0 = y_0, \dots, Y_n = y_n, Y_{n+1} = y_{n+1}) = \prod_{k=0}^{n} P(y_k, y_{k+1})$ for all $(y_0, \dots, y_{n+1}) \in \mathbb{N}^{n+2}$ and $n \in \mathbb{N}$ (the quantity is zero if $(y_0, y_1, \dots, y_{n+1})$ is not such that $y_0 = 0$ and $y_{\ell+1} \in \{0, y_\ell + 1\}$ for $\ell \in \{0, \dots, n\}$). This gives that Y is a Markov chain on $\mathbb{N}$ with transition matrix P starting from 0.

2. (a) Use that $\rho(1) = \mathbb{P}(D \geq 2)$ to deduce the result.

(b) Let $n \in \mathbb{N}^*$. We deduce from the previous question that $P^n(0, n) = \mathbb{P}(D \geq n+1)$, and this quantity is positive as D is unbounded. Since D is unbounded and takes values in $\mathbb{N}^*$, we deduce that $\{\ell \in \mathbb{N}^* : 1 - \rho(\ell) = \mathbb{P}(D = \ell \,|\, D \geq \ell) > 0\}$ is unbounded. We deduce that for all $k \in \mathbb{N}$, there is $n \geq 0$ such that $P(n + k, 0) > 0$, and thus:

$$P^{n+1}(k, 0) \geq \left(\prod_{\ell=0}^{n-1} P(k + \ell, k + \ell + 1)\right) P(n + k, 0) > 0.$$

We deduce that Y is irreducible.

3. Notice that D_1 is the first return time to 0 for Y. Since D_1 takes values in $\mathbb{N}^*$, it is finite a.s., and we get that $\mathbb{P}(D_1 < \infty) = 1$ and thus Y is recurrent. It is positive recurrent if and only if D_1 (or equivalently D) is integrable.

4. Let π be the invariant probability measure (on $\mathbb{N}^*$), which exists and is unique as Y is irreducible positive recurrent. We first notice that $\pi(0) = 1/\mathbb{E}[D]$ as D is distributed as D_1 the first return time to 0 for Y.

Denote by F the cumulative distribution of D and by $m = \mathbb{E}[D]$ its the mean. Since $\pi P = \pi$, we deduce that $\pi(n) = \rho(n)\pi(n-1)$ for all $n \in \mathbb{N}^*$, and thus, thanks to Question 2a, $\pi(n) = \pi(0) \prod_{\ell=1}^{n-1} \rho(\ell) = \mathbb{P}(D \geq n+1)/\mathbb{E}[D]$. This gives that:

$$\pi(n) = \frac{1 - F(n)}{m} \quad \text{for all} \quad n \in \mathbb{N}.$$

Exercise 2

1. Once $n \geq n_0$, there is no better strategy than to stop at the first empty place. So, if $X_{n_0} = 0$ the optimal loss at step n_0 is 0, and if $X_{n_0} = 1$, then the optimal loss is $\mathbb{E}[T|\mathcal{F}_{n_0}]$ with $T = \inf\{k \geq 1 : X_{n_0+k} = 0\}$. Since T is independent of $\mathcal{F}_{n_0}$ and is a geometric random variable with parameter q, we deduce that $\mathbb{E}[T|\mathcal{F}_{n_0}] = 1/q$ and thus $S_{n_0} = X_{n_0}/q$.

2. Notice that a loss can be seen as minus a gain. We recall the Snell envelope $(S_n, n \in \mathbb{N}^*)$ satisfies the optimal equations:

$$S_n = \min(L_n, \mathbb{E}[S_{n+1}|\mathcal{F}_n]) \quad \text{for all} \quad n \in \mathbb{N}^*.$$

We prove (1) by a backward induction. Clearly (1) is satisfied for $n = n_0$. Assume (1) holds at step $n+1$ for $n \in \{1, \dots, n_0 - 1\}$, and let us prove it holds at step n. According to the optimal equation, we have:

$$S_n = \min((n_0 - n), \mathbb{E}[S_{n+1}|\mathcal{F}_n])(1 - X_n) + X_n \mathbb{E}[S_{n+1}|\mathcal{F}_n].$$

Since (1) holds at step $n+1$, we get $\mathbb{E}[S_{n+1}|\mathcal{F}_n] = p\alpha_{n+1} + q\min(n_0 - n - 1, \alpha_{n+1}) = \alpha_n$, and thus:

$$S_n = \alpha_n X_n + \min((n_0 - n), \alpha_n)(1 - X_n).$$

This proves that (1) holds for $n \in \{1, \dots, n_0\}$.

3. Setting $L_\infty = \infty$ and taking the gain G_n equal to $-L_n$ for $n \in \bar{\mathbb{N}}^*$, we deduce that $\mathbb{E}[\sup_{n \in \bar{\mathbb{N}}^*} \max(G_n, 0)] = 0$ and a.s. $\limsup_{n \uparrow \infty} G_n = G_\infty$. We deduce from the optimal stopping theorem that the following stopping times (with the convention that $\inf \emptyset = \infty$) are optional:

$$\tau_* = \inf\{n \in \mathbb{N}^* : S_n = L_n\} \quad \text{and} \quad \tau_{**} = \inf\{n \in \mathbb{N}^* : S_n < \mathbb{E}[S_{n+1}|\mathcal{F}_n]\}. \tag{3}$$

Notice all the optimal stopping times are bounded from above by $n_0 + T$ and thus are finite a.s.. Since the loss corresponding to taking the first empty place is finite, we deduce that the optimal loss is finite, and thus a stopping time τ is optimal if and only if $\tau_* \leq \tau \leq \tau_{**}$ and $S_\tau = L_\tau$.

4. (a) As $p \leq 1/2$, we deduce that $p/q \leq 1$. By (2), we get $\alpha_{n_0-1} = p/q \leq 1$. Thus by an immediate backward induction, we deduce that $\alpha_n = p/q \leq 1$ for $n \in \{1, \dots, n_0 - 1\}$.

   (b) For all $n \in \{1, \dots, n_0 - 1\}$, we get from (1) that $S_n = \alpha_n = p/q$ as $n_0 - n \geq \alpha_n$. If $p < 1/2$, we observe that $S_n < L_n$ for all $n < n_0$, and thus $\tau_* = \inf\{n \geq n_0 : X_n = 0\}$ (that, is wait until you are in front of the opera and then take the first empty parking place). If $p = 1/2$, then we get that $S_n < L_n$ for all $n < n_0 - 1$ and thus $\tau_* = \inf\{n \geq n_0 - 1 : X_n = 0\}$ (that is, the same strategy plus the fact that you can take the parking place just before the one in front of the opera). Since $p \leq 1/2$ and $\mathbb{E}[S_{n+1}|\mathcal{F}_n] = p/q \leq 1$ for $n \in \{1, \dots, n_0 - 1\}$, we deduce that $\tau_{**} = \inf\{n \geq n_0 : X_n = 0\}$.

   In conclusion, if $p \leq 1/2$, an optimal strategy is to wait to be in front of the opera and then take the first empty parking place. Furthermore this strategy is the only optimal one if $p < 1/2$.

5. (a) If $\alpha_{n+1} \leq n_0 - n - 1$, then we deduce from (2) that $\alpha_n = \alpha_{n+1}$ and thus $\alpha_n < n_0 - n$. By contraposition, we get that $\alpha_n > n_0 - n$ implies that $\alpha_{n+1} > n_0 - n - 1$. Then use the definition of n_* to get the last part.

(b) We consider the optimal stopping time τ_* defined in (3). For $n < k_*$ and $n \geq 1$, we have $S_n = \alpha_n < L_n$. This gives that $\tau_* \geq n_*$. For $n \geq n_*$, we have $S_n = \alpha_n X_n + (n_0 - n)(1 - X_n) \leq L_n$ with an equality if and only if $X_n = 0$. This gives that $\tau_* = \inf\{n \geq n_* : \; X_n = 0\}$.

We consider the optimal stopping time τ_{**} defined in (3). We have $\tau_{**} \geq \tau_* \geq n_*$. For $n \geq n_*$, we have $S_n = \alpha_n X_n + (n_0 - n)(1 - X_n) < \alpha_n$ if and only if $X_n = 0$ and $n_0 - n < \alpha_n$. This latter condition is satisfied for all $n > n_*$. So if $\alpha_{n_*} > n_0 - n_*$, then $\tau_{**} = \tau_*$ (and thus τ_* is the only optimal stopping time). If $\alpha_{n_*} = n_0 - n_*$, then $\tau_{**} = \inf\{n > n_* : \; X_n = 0\}$, and thus $\tau_{**} > \tau_*$ on $\{X_{n_*} = 0\}$.

(c) Use that $\alpha_n = p\alpha_{n+1} + q(n_0 - n - 1)$ for $n \geq \max(n_* - 1, 1)$ to get the result by backward induction.

(d) Using that $\alpha_{n_*} \geq n_0 - n_*$ and, provided that $n_* > 1$, $\alpha_{n_*-1} < n_0 - n_* + 1$, we deduce that for $n_* = n_0 - r \in \{2, \ldots, n_0\}$ we have $p \in [2^{-1/r}, 2^{-1/(r+1)})$ and that $n_* = 1$ for $p \in [2^{-1/(n_0-1)}, 1)$.

(Notice the result is consistent with Question 4 where $n_* = 0$ if $p < 1/2$ and $n_* = 1$ if $p = 1/2$. Notice also that the stopping times τ_* and τ_{**} are not equal if and only if p is equal to $2^{-1/r}$ for some $r \in \{1, \ldots, n_0 - 1\}$.)

6. This is more delicate. One can assume that the parameter is a random $(0,1)$-valued random variable P, and that the random variable X_n is written as $\mathbf{1}_{\{U_n \leq P\}}$, where $(U_n, n \in \mathbb{N}^*)$ are independent random variables uniformly distributed on $[0,1]$ and independent of P. We still keep $\mathcal{F}_n = \sigma(X_1, \ldots, X_n)$, so that P is $\mathcal{F}_\infty$ measurable (as a.s. $P = \lim_{m\to\infty} n^{-1} \sum_{k=1}^n X_k$).

Arguing as in Question 1, we deduce that $S_{n_0} = X_{n_0} R$, where $R = \mathbb{E}[T | \, \mathcal{F}_{n_0}] = \mathbb{E}[P^{-1} | \, \mathcal{F}_{n_0}]$. We can again use the optimal equations to get $(S_n, n \in \{1, \ldots, n_0\})$, but there will be no closed formula as (1), as one as to take into account the random term R in the definition of S_n.

Let us mention that taking P following a β distribution will give a closed formula for R, but certainly not nice enough to have a simple formula for τ_* and τ_{**} (which should not belong to the family of stopping times $\inf\{k \geq k' : \; X_k = 0\}$ with k' deterministic).

Lastly, the optimal stopping times will depend on the law of P which has to be known. Informally this prior distribution of P corresponds to knowledge of previous experiences.