(*public-A1*)

Résumé : Le gérant d'un casino s'intéresse à l'évolution de son capital en fonction du temps. Il souhaite en particulier étudier la probabilité qu'il puisse un jour être ruiné.

Mots clefs : Loi de Poisson, lois exponentielles, marches aléatoires, conditionnement, théorèmes asymptotiques.

- *Il est rappelé que le jury n'exige pas une compréhension exhaustive du texte. Il vous est conseillé de construire un exposé évitant la paraphrase et mettant en lumière vos connaissances, à partir des éléments du texte. Vous êtes libre d'organiser votre discussion comme vous l'entendez. Des pistes de réflexion, largement indépendantes les unes des autres, vous sont proposées en fin de texte; vous n'êtes pas tenu de les suivre. Le propos devra être illustré par des traitements ou des simulations numériques sur ordinateur, ou, à défaut, des propositions de telles illustrations. Le jury souhaiterait que le plan de la présentation soit annoncé au début de l'exposé.*

Dans une première partie seront présentés le modèle et la construction d'un processus aléatoire $(Y_t)_{t\geqslant 0}$ représentant le capital du casino en fonction du temps. Dans une deuxième partie, il sera établi que si le gérant choisit mal le paramètre caractérisant les rentrées d'argent (en faisant payer les jetons de jeu trop peu cher ou en n'incitant pas suffisamment les clients à jouer par exemple), alors il sera presque sûrement ruiné un jour, et ce quel que soit son investissement initial! Dans une troisième partie, le texte présentera comment rendre la probabilité d'être ruiné exponentiellement petite (dans le cas bien sûr où le précédent paramètre est bien choisi). Enfin dans une quatrième partie, on verra qu'on peut calculer explicitement cette probabilité si la distribution des gains des joueurs est exponentielle.

1. Le modèle

Les rentrées d'argent du casino (c'est-à-dire les dépenses des clients) sont déterministes et gouvernées par un paramètre $\alpha > 0$ de la façon suivante : entre les instants 0 et t, le casino gagne αt.

Les gains des joueurs ont lieu aux instants de saut d'un processus de Poisson $(N_t)_{t\geqslant 0}$ de paramètre 1. Plus précisément, considérons $(\xi_k)_{k\geqslant 1}$ une suite de variables aléatoires indépendantes identiquement distribuées (ce que l'on notera i.i.d. dans la suite) de loi exponentielle de paramètre 1 et posons : $T_0 = 0$,

$$\forall i \in \mathbb{N}^*,\ T_i = \xi_1 + \cdots + \xi_i\,; \quad \text{et} \quad \forall t \in \mathbb{R}_+,\ N_t = \max\{i \in \mathbb{N} : T_i \leqslant t\}.$$

Ces gains sont donnés par une suite de variables aléatoires strictement positives $(X_i)_{i\geqslant 1}$ i.i.d, et indépendantes du processus $(N_t)_{t\geqslant 0}$. Ainsi, à l'instant T_1, un premier joueur gagne X_1, puis à l'instant T_2, un deuxième joueur gagne X_2, etc. Remarquons que la variable N_t est égale au nombre de joueurs ayant eu un gain avant l'instant t.

On suppose que X_1 admet un moment d'ordre 2. On note

$$\mu = \mathbb{E}(X_1) \quad \text{et} \quad \sigma^2 = \mathrm{Var}(X_1).$$

Pour tout $t \in \mathbb{R}_+$, notons Y_t le capital du casino à l'instant t. On a alors, si $Y_0 \geqslant 0$ est le capital initial du casino, pour tout instant $t \geqslant 0$,

$$Y_t = Y_0 + \alpha t - \sum_{i=1}^{N_t} X_i.$$

(Lorsque $N_t = 0$ la somme de cette formule est par définition égale à 0.)

L'objectif principal du gérant est d'évaluer en fonction de son investissement initial y, la probabilité $r(y)$ qu'il ait un jour tout perdu, soit

$$r(y) = \mathbb{P}(\exists t \in \mathbb{R}_+ \text{ t.q. } Y_t < 0 \mid Y_0 = y).$$

Bien évidemment, de son point de vue, l'idéal est que cette probabilité soit la plus petite possible. Malheureusement pour lui,

Proposition 1. $\forall y \geqslant 0$, $r(y) > 0$.

DÉMONSTRATION – Considérons deux réels $t > 0$ et $y \geqslant 0$. On peut trouver un entier $k \geqslant 1$ tel que

$$\mathbb{P}(X_1 + \cdots + X_k > y + \alpha t) > 0.$$

Comme par ailleurs $\mathbb{P}(N_t = k) > 0$, on a $\mathbb{P}(Y_t < 0 \mid Y_0 = y) > 0$, ce qui amène au résultat. □

La probabilité $r(y)$ est donc strictement positive quel que soit l'investissement initial y. Le but du gérant est donc de choisir les paramètres de son modèle pour minimiser cette probabilité.

2. Choix du paramètre α

Proposition 2. *Si $\alpha < \mu$ alors : $\forall y \geqslant 0$, $r(y) = 1$.*

DÉMONSTRATION – Pour tout $t \geqslant 0$, on a $N_t = (N_t - N_{\lfloor t \rfloor}) + \sum_{i=1}^{\lfloor t \rfloor} (N_i - N_{i-1})$ (où $\lfloor t \rfloor$ désigne la partie entière de t). Un résultat classique sur les processus de Poisson (on pourra admettre ce résultat au besoin) dit que les variables $(N_i - N_{i-1})$ forment une suite de variables aléatoires i.i.d. de loi de Poisson de paramètre 1. Ainsi d'après la loi forte des grands nombres, on a : presque sûrement (p.s.), quand $t \to +\infty$,

$$\frac{N_t}{t} \longrightarrow 1.$$

En utilisant à nouveau la loi forte des grands nombres et par un argument de composition de limites, on obtient alors : p.s., quand $t \to +\infty$,

$$\frac{Y_t}{t} \longrightarrow \alpha - \mu,$$

ce qui nous amène au résultat. □

Cette démonstration ne permet pas de conclure lorsque $\alpha = \mu$, pourtant le résultat reste vrai dans ce cas.

Proposition 3. *Si $\alpha = \mu$ alors : $\forall y \geqslant 0$, $r(y) = 1$.*

Pour montrer ce résultat, on va introduire la marche aléatoire définie pour $n \geqslant 1$ par

$$S_n = \sum_{i=1}^{n} (X_i - \alpha \xi_i). \tag{1}$$

DÉMONSTRATION – On a : $Y_{T_n} = y - S_n$. Pour prouver le résultat, il suffit par exemple de vérifier que $\limsup_{n\to\infty} S_n = \infty$ presque sûrement lorsque $\alpha = \mu$. Pour cela, remarquons que si $\alpha = \mu$,

$$\begin{aligned}
\mathbb{P}\left(\limsup \frac{S_n}{\sqrt{n}} \geqslant 1\right) \geqslant \mathbb{P}\left(\limsup \left\{\frac{S_n}{\sqrt{n}} > 1\right\}\right) &= \lim_{n\to+\infty} \mathbb{P}\left(\bigcup_{k\geqslant n} \left\{\frac{S_k}{\sqrt{k}} > 1\right\}\right) \\
&\geqslant \lim_{n\to+\infty} \mathbb{P}\left(\frac{S_n}{\sqrt{n}} > 1\right) \\
&> 0.
\end{aligned}$$

Or, pour tout N entier, l'événement $\limsup\{S_n/\sqrt{n} > 1\}$ ne dépend pas des variables $X_1 - \alpha\xi_1, \ldots, X_N - \alpha\xi_N$. Ainsi il appartient à la tribu asymptotique de la suite $(X_i - \alpha\xi_i)_{i\geqslant 1}$ qui est triviale d'après la loi du 0-1 de Kolmogorov (rappelée ci-dessous). D'où,

$$\mathbb{P}\left(\limsup \frac{S_n}{\sqrt{n}} \geqslant 1\right) = 1.$$

□

Lemme 1. Loi du 0-1 de Kolmogorov. *Soit $(Z_n)_{n\geqslant 1}$ une suite de variables aléatoires indépendantes. On définit la tribu asymptotique $\mathscr{A}$ de $(Z_n)_{n\geqslant 1}$ par*

$$\mathscr{A} = \bigcap_{n\geqslant 1} \sigma(Z_n, Z_{n+1}, \ldots).$$

Alors la tribu $\mathscr{A}$ est triviale i.e., pour tout événement E de $\mathscr{A}$, $\mathbb{P}(E) \in \{0,1\}$.

Le gérant du casino a donc intérêt à choisir un coefficient $\alpha > \mu$.

3. Majoration de la probabilité $r(y)$ dans le cas $\alpha > \mu$

Afin d'obtenir des renseignements sur la probabilité $r(y)$ lorsque $\alpha > \mu$, on fait l'hypothèse suivante sur le modèle :

(H_A) $\exists A > 0$ tel que e^{AX_1} soit intégrable et

$$\mathbb{E}\left(e^{A(X_1 - \alpha T_1)}\right) = 1. \tag{2}$$

Notons que ceci définit de façon non ambiguë un coefficient A relié au modèle, puisqu'on a le résultat d'unicité suivant.

Proposition 4. *L'équation (2) admet au plus une solution strictement positive.*

DÉMONSTRATION – Il suffit de remarquer que la fonction $x \mapsto \mathbb{E}\left(e^{x(X_1-\alpha T_1)}\right)$ est strictement convexe sur son domaine de définition. □

Par exemple, si X_1 suit une loi exponentielle de paramètre $\gamma > 1/\alpha$, $A = \gamma - 1/\alpha$. De façon générale, l'intérêt de l'existence de ce coefficient A se voit dans le résultat suivant.

Proposition 5. *Sous l'hypothèse* (H_A)*, la probabilité de ruine vérifie* $r(y) \leqslant e^{-Ay}$, $\forall y \geqslant 0$.

DÉMONSTRATION – On utilise la marche aléatoire $(S_n)_{n\geqslant 1}$ définie en (1). Notons qu'il suffit de se placer aux instants de saut pour pouvoir observer la ruine, d'où :

$$r(y) = \mathbb{P}\left(\max_{n\geqslant 1} S_n > y\right) = \lim_{n\to+\infty} \mathbb{P}\left(\max_{1\leqslant k\leqslant n} S_k > y\right).$$

Pour démontrer la proposition, il suffit donc d'avoir $r_n(y) = \mathbb{P}\left(\max_{1\leqslant k\leqslant n} S_k > y\right) \leqslant e^{-Ay}$ pour tout entier n et pour tout $y \geqslant 0$.

On a bien $r_1(y) \leqslant e^{-Ay}$, $\forall y \geqslant 0$, d'après l'inégalité de Markov. Supposons donc, pour un entier n, que $r_n(y) \leqslant e^{-Ay}$, $\forall y \geqslant 0$. Soit alors $y \geqslant 0$ fixé. On a

$$r_{n+1}(y) = \mathbb{P}(S_1 > y) + \mathbb{P}\left(S_1 \leqslant y, \max_{2\leqslant k\leqslant n+1}(S_k - S_1) > y - S_1\right).$$

D'une part, notant $\mathbb{P}_{S_1}$ la loi de S_1, on a

$$\mathbb{P}(S_1 > y) = \int_{]y,\infty[} \mathbb{P}_{S_1}(\mathrm{d}u) \leqslant \int_{]y,\infty[} e^{-A(y-u)}\mathbb{P}_{S_1}(\mathrm{d}u).$$

D'autre part,

$$\mathbb{P}\left(S_1 \leqslant y, \max_{2\leqslant k\leqslant n+1}(S_k - S_1) > y - S_1\right) = \int_{[0,y]} r_n(y-u)\mathbb{P}_{S_1}(\mathrm{d}u) \leqslant \int_{[0,y]} e^{-A(y-u)}\mathbb{P}_{S_1}(\mathrm{d}u).$$

On obtient alors $r_{n+1}(y) \leqslant e^{-Ay}$, $\forall y \geqslant 0$, ce qui achève la récurrence et la preuve de cette proposition. □

4. Calcul exact : cas des gains de loi exponentielle

La probabilité $r(y)$ est difficile à calculer en général. Il y a cependant un cas où on peut l'exprimer explicitement : lorsque les gains $(X_i)_{i\geqslant 1}$ sont distribués suivant une loi exponentielle, disons de paramètre $\gamma = 1/\mu > 1/\alpha$.

Théorème 1. *Dans ce cas,*

$$r(y) = \frac{e^{-(\gamma - 1/\alpha)y}}{\alpha\gamma}, \quad \forall y \geqslant 0.$$

DÉMONSTRATION – On travaille à nouveau avec la marche aléatoire $(S_n)_{n\geqslant 1}$. Introduisons

$$g(y) := 1 - r(y) = \mathbb{P}(\max_{n\geqslant 1} S_n \leqslant y),$$

la probabilité que le capital du casino soit toujours positif. On a alors

$$\begin{aligned} g(y) &= \mathbb{P}\left(S_1 \leqslant y, \max_{n\geqslant 2}(S_n - S_1) \leqslant y - S_1\right) \\ &= \int_0^\infty \left(\int_0^{y+\alpha u} g(y - x + \alpha u)\gamma e^{-\gamma x}\mathrm{d}x\right) e^{-u}\mathrm{d}u \\ &= \frac{e^{y/\alpha}}{\alpha}\int_y^\infty \left(\int_0^z g(v)\gamma e^{\gamma v}\mathrm{d}v\right) e^{-(\gamma+1/\alpha)z}\mathrm{d}z. \end{aligned}$$

En dérivant, on obtient

$$g'(y) = \frac{1}{\alpha}\left(g(y) - e^{-\gamma y}\int_0^y g(v)\gamma e^{\gamma v}\mathrm{d}v\right),$$

qui nous amène à

$$r'(y) = \frac{1}{\alpha}\left(r(y) - e^{-\gamma y} - \gamma e^{-\gamma y}\int_0^y r(v)e^{\gamma v}\mathrm{d}v\right).$$

Enfin, posons $h(y) = \int_0^y r(v)e^{\gamma v}\mathrm{d}v$. Cette fonction est solution de l'équation différentielle du second ordre

$$h''(y) = \left(\gamma + \frac{1}{\alpha}\right)h'(y) - \frac{\gamma}{\alpha}h(y) - \frac{1}{\alpha},$$

d'où le résultat. $\square$

Le gérant du casino a donc un intérêt certain à savoir si les gains des joueurs sont distribués exponentiellement ou pas. Pour en avoir le cœur net, il décide de mettre en place un test pour évaluer l'ajustement de l'échantillon $(X_1, ..., X_n)$ à la famille des lois exponentielles. Il procède de la façon suivante : un estimateur du paramètre d'une loi exponentielle étant

$$\hat{\gamma}_n = \frac{1}{\overline{X}_n} := \frac{n}{X_1 + ... + X_n},$$

il décide d'introduire la statistique

$$H_n = \sup_{x\geqslant 0}\left|\hat{F}_n(x) - (1 - e^{-x/\overline{X}_n})\right|,$$

où $\hat{F}_n$ désigne la fonction de répartition empirique de l'échantillon $(X_1, ..., X_n)$.

On remarque dans un premier temps que si les $(X_i)_{i\geqslant 1}$ ont une distribution exponentielle, $H_n \to 0$ presque sûrement quand $n \to \infty$, tandis que H_n aura une limite presque sûre strictement positive si la distribution n'est pas exponentielle. Dans la mise en place d'un test pour évaluer l'adéquation à la famille des lois exponentielles, il est donc naturel d'introduire une zone de rejet de la forme $\{H_n \geqslant C\}$ avec une constante C adéquate. Reste ensuite à déterminer cette constante. Il se trouve qu'on a le résultat suivant.

Lemme 2. *Si les $(X_i)_{i\geqslant 1}$ ont une distribution exponentielle de paramètre γ, la loi de la statistique H_n est indépendante de γ.*

DÉMONSTRATION – Posons $X'_i = \gamma X_i$. Alors, sous l'hypothèse de l'énoncé, $(X'_1, ..., X'_n)$ est un échantillon de loi exponentielle de paramètre 1. Notant F'_n sa fonction de répartition empirique, on voit que H_n peut se réécrire

$$H_n = \sup_{x \geqslant 0} \left| \hat{F}'_n(x) - (1 - e^{-x/\overline{X}'_n}) \right|,$$

d'où la conclusion. □

Fort de ce résultat, le gérant n'a plus qu'à simuler le quantile adéquat de la loi de H_n sous l'hypothèse exponentielle pour pouvoir mettre en place son test. Si l'hypothèse d'exponentialité est acceptée, il pourra ensuite utiliser un estimateur du paramètre pour avoir une idée du comportement de la probabilité $r(y)$.

Suggestions et pistes de réflexion

► *Les pistes de réflexion suivantes ne sont qu'indicatives : vous n'êtes pas obligé de les suivre. Vous pouvez choisir d'étudier certains des points proposés, de façon plus ou moins approfondie, mais aussi toute autre question à votre initiative. Vos investigations comporteront une partie traitée sur ordinateur et, si possible, des représentations graphiques de vos résultats. À défaut, si vos traitements ou simulations numériques n'ont pas abouti, il est conseillé d'expliquer ce que vous auriez souhaité mettre en œuvre.*

- Aspects mathématiques :
  - On pourra compléter certaines preuves des différentes propositions, lemmes et théorème du texte.
  - On pourra montrer les assertions sur la statistique H_n introduite dans le dernier paragraphe.
- Simulations :
  - On pourra représenter graphiquement le processus $(Y_t)_{t \geqslant 0}$.
  - On pourra illustrer numériquement les résultats des Propositions 1, 2, 3 et du Théorème 1. À cet effet, comment calcule-t-on une approximation de $r(y)$?
- Modélisation :
  - On pourra discuter la pertinence du modèle, par exemple discuter les gains déterministes αt, la distribution des lois des temps d'inter-arrivées, etc. Pourquoi n'observe-t-on jamais de ruine en pratique ? Il y a pourtant beaucoup de casinos !