Sujet 1 - Probabilités et Statistiques

Préliminaire

Le but de ce problème est de montrer certains résultats en théorie de la ruine, théorie utilisée notamment pour des applications à l'étude de la solvabilité d'une compagnie d'assurance. On considère la suite de variables aléatoires réelles $(C_n(u,p))_{n\in\mathbb{N}}$ telle que, pour $n > 0$:

$$C_n(u,p) = u + pn - \sum_{i=1}^{n} S_i,$$

et $C_0(u,p) = u$. Dans cette relation, $u > 0$ est une fortune initiale (déterministe), $p > 0$ est une quantité (déterministe) de primes reçues par la compagnie d'assurance entre deux dates, et S_i correspond aux montants (aléatoires) des règlements de sinistres effectués entre la date $i-1$ et la date i. $C_n(u,p)$ représente alors la fortune disponible à la date n. Le but de ce problème consiste à évaluer le mieux possible la probabilité de ruine, i.e.

$$\pi(u,p) = \mathbb{P}\left(\exists n \geq 0 : C_n(u,p) < 0\right),$$

en fonction de différentes hypothèses sur u, p, et les $(S_i)_{i\geq 1}$.

Dans toute la suite de ce problème, on se place dans un espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$, où $\mathcal{F}$ désigne une tribu sur l'ensemble Ω, et $\mathbb{P}$ est une mesure de probabilité sur $\mathcal{F}$. De plus, E et Var désignent respectivement la moyenne et la variance pour la probabilité $\mathbb{P}$.

Rappels

Lois de Poisson et exponentielle : Soit $\lambda > 0$. Une variable aléatoire Z_1 suit une loi de Poisson si, pour tout $k \in \mathbb{N}$,

$$\mathbb{P}(Z_1 = k) = e^{-\lambda}\frac{\lambda^k}{k!}.$$

Une variable aléatoire Z_2 suit une loi exponentielle de paramètre λ si, pour tout borélien A de $\mathbb{R}$,

$$\mathbb{P}(Z_2 \in A) = \int_A \lambda e^{-\lambda x}\mathrm{d}x.$$

Filtrations : On rappelle qu'une filtration $(\mathcal{F}_n)_{n\in\mathbb{N}}$ est une suite de sous-tribus de $\mathcal{F}$ telle que

$$\forall n \in \mathbb{N},\ \mathcal{F}_n \subset \mathcal{F}_{n+1}.$$

Martingales : Soit $(\mathcal{F}_n)_{n\in\mathbb{N}}$ une filtration. Une suite de variables aléatoires réelles $(X_n)_{n\in\mathbb{N}}$ est une $(\mathcal{F}_n)_{n\in\mathbb{N}}-$martingale si, par définition, les trois conditions suivantes sont vérifiées :

i) $\forall n \in \mathbb{N},\ X_n$ est $\mathcal{F}_n-$mesurable,

ii) $\forall n \in \mathbb{N},\ E[|X_n|] < \infty$,

iii) $\forall n \in \mathbb{N},\ E[X_{n+1}|\mathcal{F}_n] = X_n$,

où $E[X|\mathcal{B}]$ désigne l'espérance conditionnelle d'une variable aléatoire réelle et intégrable X par rapport à $\mathcal{B}$, sous-tribu de $\mathcal{F}$.

Temps d'arrêt :

Soit $(\mathcal{F}_n)_{n\in\mathbb{N}}$ une filtration. On appelle $(\mathcal{F}_n)_{n\in\mathbb{N}}-$temps d'arrêt une variable aléatoire T telle que

i) T prend ses valeurs dans $\mathbb{N}\cup\{+\infty\}$ presque sûrement,

ii) $\forall n \in \mathbb{N},\ \{T = n\} \in \mathcal{F}_n$.

Convergence en loi :

On rappelle qu'une suite de vecteurs aléatoires $(Z_n)_{n\in\mathbb{N}}$ à valeurs dans $\mathbb{R}^p$ converge en loi vers un vecteur aléatoire Z si, pour toute fonction $f : \mathbb{R}^p \to \mathbb{R}$ telle que f soit continue et bornée,

$$E[f(Z_n)] \to_{n\to\infty} E[f(Z)].$$

1 Utilisation de la théorie des martingales

Dans cette partie, on utilise la théorie des martingales pour obtenir une majoration de la probabilité de ruine dans le cas où les S_i sont indépendantes, identiquement distribuées (i.i.d.) et de même loi qu'une variable aléatoire $S = \sum_{j=1}^{N} X_j$ (avec la convention $\sum_{j=1}^{N} X_j = 0$ lorsque $N = 0$). Dans le cadre de la modélisation du capital d'une compagnie d'assurance, N est un nombre aléatoire de sinistres se produisant dans une période de temps, et X_j est le coût aléatoire du $j-$ème sinistre. On suppose que les $(X_j)_{j\geq 1}$ sont des variables aléatoires réelles presque sûrement positives.

1.1 Soit $(M_n)_{n\in\mathbb{N}}$ une $(\mathcal{F}_n)_{n\in\mathbb{N}}-$martingale, et soit T un $(\mathcal{F}_n)_{n\in\mathbb{N}}-$temps d'arrêt. On suppose que $T \leq \tau$ p.s. où τ est un réel. Montrer que

$$E[M_T] = E[M_0].$$

Indication: On pourra utiliser l'identité $M_T = M_0 + \sum_{i=1}^{\tau}(M_i - M_{i-1})\mathbf{1}_{\{T\geq i\}}$.

1.2 On suppose de plus que, pour tout $n \geq 0$, $M_n \geq 0$ p.s. On définit, pour $x \in \mathbb{R}^+$, $A_m(x) = \{\sup_{0\leq k\leq m} M_k > x\}$, et

$$T_m(x) = \min\{k : 0 \leq k \leq m, M_k > x\}\mathbf{1}_{A_m(x)} + m\mathbf{1}_{A^c_m(x)},$$

où A^c désigne le complémentaire d'un ensemble A. Montrer que $T_m(x)$ est un $(\mathcal{F}_n)_{n\in\mathbb{N}}-$temps d'arrêt.

1.3 Sous les hypothèses de la question 1.2, montrer que $x\mathbb{P}(A_m(x)) \leq E[M_0]$.

1.4 Dans la suite de cette partie, N est une variable aléatoire de loi de Poisson de paramètre $\lambda > 0$, et $(X_i)_{i\geq 1}$ est une suite de variables aléatoires i.i.d. indépendantes de la variable aléatoire N. Montrer que $\mathbb{P}(S < +\infty) = 1$.

1.5 On suppose que $\mu_1 = E[X_1] < \infty$. Déterminer $E[S]$.

1.6 On suppose que $\mu_2 = E[X_1^2] < \infty$. Déterminer $Var(S)$.

1.7 Dans la suite de cette partie, on suppose qu'il existe $a > 0$ tel que $\phi(a) = E[\exp(aX_1)] < \infty$. Calculer

$$\psi(a) = E[\exp(aS)].$$

1.8 On considère des variables aléatoires $(S_i)_{i\geq 1}$ i.i.d. de même loi que S. Déterminer $s(a)$ pour que la suite de variables aléatoires $(M_n)_{n\in\mathbb{N}}$ définie par

$$\begin{aligned} M_0 &= 1, \\ M_n &= \exp\left(a\sum_{i=1}^{n} S_i - ns(a)\right) \text{ si } n > 0, \end{aligned}$$

soit une $(\mathcal{F}_n)_{n\in\mathbb{N}}-$martingale, où $\mathcal{F}_0 = \{\emptyset, \Omega\}$ et, pour $n > 0$, $\mathcal{F}_n$ est la tribu engendrée par $\{S_1, ..., S_n\}$.

1.9 Soit $x > 0$. En déduire une majoration de $\mathbb{P}(\sup_{1\leq k\leq m}\{\sum_{i=1}^{k} S_i - ks(a)/a\} > x)$.

1.10 Expliciter $s(a)$ dans le cas où les X_i sont des variables exponentielles de paramètre $\mu > 0$ (i.e. d'espérance $1/\mu$), en précisant l'ensemble de définition.

1.11 En déduire, toujours dans le cas où les X_i sont des variables exponentielles de paramètre $\mu > 0$, une borne supérieure pour p_0 défini par $p_0 = \inf(p : \pi(1,p) \leq \exp(-a))$, en supposant $\mu > a$.

2 Borne de Lundberg

Dans toute cette partie, les S_i sont i.i.d. de même loi que

$$S = \sum_{j=1}^{N} X_j$$

(avec la convention $\sum_{j=1}^{N} X_j = 0$ lorsque $N = 0$), où les $(X_j)_{j\geq 1}$ sont des variables aléatoires positives i.i.d. indépendantes de N. On suppose également que N suit une loi de Poisson de paramètre $\lambda > 0$. On note

$$\begin{aligned}
\mu_1 &= E[X_1], \text{ supposé non nul,} \\
\phi(r) &= E[\exp(rX_1)], \text{ éventuellement infini,} \\
\mathcal{R} &= \{r \in \mathbb{R} : \phi(r) < \infty\}.
\end{aligned}$$

Soit R la plus grande solution positive appartenant à $\mathcal{R}$ de l'équation d'inconnue r suivante :

$$1 + \frac{p}{\lambda} r = \phi(r). \tag{1}$$

2.1 Soit X_1 une variable aléatoire telle que $\mathcal{R} =]-\infty; r_0[$, avec $r_0 > 0$. Montrer que $R = 0$ si $p \leq \lambda\mu_1$.

2.2 On suppose que sur $[0, r_0]$ avec $r_0 > 0$, ϕ est définie et de plus $E[X_1^2 \exp(r_0 X_1)] < \infty$. Montrer que ϕ est strictement convexe sur $[0, r_0]$.

2.3 En déduire que, sous les hypothèses de la question précédente, l'équation (1) admet au plus une solution sur $]0, r_0]$.

2.4 On suppose à présent que, pour $r_1 > 0$ et pour tout $r_0 < r_1$, $E[X_1^2 \exp(r_0 X_1)] < \infty$, et $E[\exp(r_1 X_1)] = \infty$. Montrer que, si $p > \lambda\mu_1$, l'équation (1) admet une unique solution strictement positive.

2.5 Déterminer R dans le cas où X_1 suit une loi exponentielle de paramètre $\mu > 0$ (i.e. d'espérance $1/\mu$), en fonction de p, λ et μ.

2.6 On considère à présent une variable aléatoire X_1 pour laquelle $R > 0$. Soit

$$\pi_k(u, p) = \mathbb{P}(\exists i \in \{1, ..., k\} : \ C_i(u, p) < 0).$$

Déterminer, pour $k > 0$ et $t > 0$,

$$\mathbb{P}(\exists i \in \{1, \cdots, k\} \ : \ C_i(u, p) < 0 | S_1 = t)$$

en fonction de $\pi_{k-1}(u', p)$, pour un certain u' que l'on précisera.

2.7 En déduire une démonstration par récurrence du fait que $\pi_k(u, p) \leq \exp(-Ru)$ pour tout $k \geq 1$, puis une majoration de $\pi(u, p)$.

3 Estimation des paramètres

On considère S tel que défini dans la section précédente. On suppose dans toute cette partie que N suit une loi de Poisson de paramètre $\lambda > 0$ et que X_1 suit une loi exponentielle de paramètre $\mu > 0$ avec $p > \lambda/\mu$. On rappelle que p est fixé.

Dans cette partie, $(S_1, N_1), (S_2, N_2), \cdots$ sont des vecteurs aléatoires i.i.d. de même loi que (S, N).

3.1 *On suppose dans un premier temps et jusqu'à la question 3.5 incluse que la valeur du paramètre λ est connue.* A partir de l'expression de $E[S]$, construire un estimateur $\hat{\mu}_n$ de μ, i.e. une fonction borélienne de $(S_1, \cdots, S_n)$, qui converge presque sûrement vers μ lorsque $n \to \infty$.

3.2 Montrer que si une suite de vecteurs aléatoires $(U_n)_{n\in\mathbb{N}}$ à valeurs dans $\mathbb{R}^d$ converge en loi vers une constante a, alors elle converge en probabilité vers a.

3.3 Soient $\alpha \in \mathbb{R}^2$ et $\varphi : \mathbb{R}^2 \to \mathbb{R}$ une fonction de classe C^1 sur un ouvert contenant α. On note $\varphi'(\alpha)$ le vecteur ligne $\varphi'(\alpha) = (\frac{\partial\varphi}{\partial x}(\alpha), \frac{\partial\varphi}{\partial y}(\alpha))$ et on considère une suite de vecteurs aléatoires (colonnes) $(Z_n)_{n\in\mathbb{N}}$ à valeurs dans $\mathbb{R}^2$, telle que $(n^{1/2}(Z_n - \alpha))_{n\in\mathbb{N}}$ converge en loi vers un vecteur aléatoire Z. Montrer que $(n^{1/2}(\varphi(Z_n) - \varphi(\alpha)))_{n\in\mathbb{N}}$ converge en loi vers $\varphi'(\alpha)Z$.

Indication: on pourra utiliser le lemme de Slutsky pour des variables aléatoires réelles, i.e. si $(A_n)_{n\in\mathbb{N}}$ converge en loi vers A, et si $(B_n)_{n\in\mathbb{N}}$ et $(C_n)_{n\in\mathbb{N}}$ convergent en loi vers des constantes b et c, alors $(A_nB_n + C_n)_{n\in\mathbb{N}}$ converge en loi vers $Ab + c$.

3.4 Calculer la loi limite de $(n^{1/2}(\hat{\mu}_n - \mu))_{n\geq 1}$.

3.5 Construire un estimateur $\hat{R}_n$ du coefficient R, i.e. une fonction borélienne de $(S_1, \cdots, S_n)$, tel que $\hat{R}_n \to R$ p.s. lorsque $n \to \infty$. Quelle est la loi limite de $(n^{1/2}(\hat{R}_n - R))_{n\geq 1}$?

3.6 *On suppose dans la suite de cette partie que λ n'est pas connu.* Construire un estimateur $\hat{\lambda}_n$ de λ, défini comme une fonction borélienne de $(N_{n+1}, \cdots, N_{2n})$, qui converge presque sûrement vers λ lorsque $n \to \infty$. Quelle est la loi limite de $(n^{1/2}(\hat{\lambda}_n - \lambda))_{n\geq 1}$?

3.7 A partir des expressions de $E[S]$ et $\hat{\lambda}_n$, construire un estimateur de $\tilde{\mu}_n$ de μ, défini comme une fonction borélienne de $(S_1, \cdots, S_n, N_{n+1}, \cdots, N_{2n})$, qui converge presque sûrement vers μ lorsque $n \to \infty$. Calculer la loi limite de $(n^{1/2}(\tilde{\mu}_n - \mu))_{n\geq 1}$.

4 Une expression de la probabilité de ruine

Le but de cette partie est de déterminer une expression satisfaite par la probabilité de ruine, et d'en déduire quelques propriétés simples sur celles-ci. On reprend les hypothèses et notations de la partie 3, avec $p > \lambda/\mu$.

4.1 Montrer qu'il existe une suite $(c_n)_{n\geq 1}$ tendant vers $+\infty$ telle que $\mathbb{P}(C_n(u,p) \leq c_n) \to_{n\to\infty} 0$.

Indication : on pourra utiliser l'inégalité de Bienaymé-Tchebychev.

4.2 Soit $T = \inf\{n \geq 0 : C_n(u,p) < 0\}$, avec la convention $\inf\{\emptyset\} = +\infty$. Montrer que, lorsque $n \to \infty$, $E[\exp(-RC_n(u,p))|T > n]\mathbb{P}(T > n)$ tend vers 0.

Indication : on pourra commencer par écrire

$$\begin{aligned} E[\exp(-RC_n(u,p))|T > n] &= E[\exp(-RC_n(u,p))\mathbf{1}_{\{C_n(u,p)\leq c_n\}}|T > n] \\ &\quad + E[\exp(-RC_n(u,p))\mathbf{1}_{\{C_n(u,p)> c_n\}}|T > n]. \end{aligned}$$

4.3 Montrer que, pour $k < n$, $E[\exp(-RC_n(u,p))|T = k] = E[\exp(-RC_T(u,p))|T = k]$.

4.4 On suppose que $\mathbb{P}(T < +\infty) > 0$. En utilisant l'identité

$$E\left[\exp(-RC_n(u,p))\right] = E[\exp(-RC_n(u,p))|T \leq n]\mathbb{P}(T \leq n) + E[\exp(-RC_n(u,p))|T > n]\mathbb{P}(T > n),$$

justifier que

$$\pi(u,p) = \frac{\exp(-Ru)}{E[\exp(-RC_T(u,p))|T < \infty]}.$$

4.5 En déduire une autre preuve de la borne de Lundberg obtenue à la question 2.7.

4.6 On suppose que $E[S_1] > p$ et qu'il existe $\alpha > 0$ tel que $S_1 \leq \alpha$ p.s. Montrer que $\mathbb{P}(T < +\infty) > 0$ et que $E[\exp(-RC_T(u,p))|T < \infty] < \infty$. En déduire également que $\pi(u,p) \to 1$ quand R tend vers 0.

5 Distribution du capital à la ruine

Dans cette partie, $S_1, S_2, \cdots$ sont des variables aléatoires i.i.d., $T_p(u) = \inf\{n \geq 0 : C_n(u,p) < 0\}$, où $\inf\{\emptyset\} = +\infty$ et, pour $y > 0$:

$$H(u,y) = \mathbb{P}(C_{T_p(u)}(u,p) \leq -y, T_p(u) < \infty).$$

Le but de cette partie est de déterminer une équation satisfaite par cette fonction H. La connaissance de cette fonction permettant, in fine, de calculer le dénominateur présent dans la question 4.4.

On considère le cas où $\mathbb{P}(S_1 = 0) = \exp(-\lambda)$ pour un paramètre $\lambda > 0$, et on désigne par $L(t)$ la fonction de survie de la loi conditionnelle de S_1 sachant $S_1 > 0$, i.e. pour tout $t \geq 0$:

$$L(t) = \mathbb{P}(S_1 > t | S_1 > 0).$$

La fonction L est supposée dérivable, et sa dérivée est notée $-l$.

5.1 On suppose pour cette question seulement que la fortune initiale u est strictement négative. Déterminer $H(u,y)$ pour $y > 0$.

5.2 Dans la suite, $u > 0$ et $y > 0$. Etablir la relation

$$\mathbb{P}(C_{T_p(u)}(u,p) \leq -y, T_p(u) = 1) = L(u+p+y)(1 - e^{-\lambda}).$$

5.3 Montrer que

$$\mathbb{P}(C_{T_p(u)}(u,p) \leq -y, S_1 = 0, T_p(u) < \infty) = H(u+p,p)e^{-\lambda}.$$

5.4 Prouver que

$$\mathbb{P}(C_{T_p(u)}(u,p) \leq -y, S_1 > 0, 1 < T_p(u) < \infty) = \int_0^x f(t)l(t)dt,$$

où on explicitera x et la valeur de f en fonction de H. En déduire que H est solution d'une équation du type

$$H(u,y) = g_{p,\lambda}(H,u,y),$$

où l'on précisera la fonction $g_{p,\lambda}$.

Fin du sujet 1