

## Processus de Poisson composés, transformation de Esscher, splitting et MCMC

### Quelques rappels en python

**random.exponential:** du module `numpy`. Pour la simulation de v.a. exponentielles. Attention à la convention sur les paramètres.  
**random.poisson:** du module `numpy`. Pour la simulation de v.a. de Poisson.  
**random.standard\_cauchy:** du module `numpy`. Pour la simulation de v.a. de Cauchy standard.  
**sort:** du module `numpy`. Pour ordonner les éléments d'un tableau.  
**stem:** du module `matplotlib.pyplot`. Pour tracer des barres verticales.  
**boxplot:** du module `matplotlib.pyplot`. Pour tracer des “boîtes à moustache” (boxplot).

*Représentation d'un array numpy par un boxplot: affichage d'une boîte rectangulaire délimitée par les premiers et derniers quartiles; la barre (rouge) du milieu donne la médiane; les “moustaches” vont jusqu'à la valeur la plus extrême dans la limite de 1.5 fois la hauteur de la boîte; les points de l'échantillon au-delà des moustaches sont représentés par des +.*

### 1. PROCESSUS DE POISSON HOMOGÈNES ET COMPOSÉS

**1.1. Rappels : Processus de Poisson.** Soit  $\lambda > 0$ . Un *processus de Poisson homogène de paramètre  $\lambda$*  est un processus  $(N_t)_{t \geq 0}$  issu de 0, à accroissement indépendants, tel que pour tout  $0 \leq s < t$ ,  $N_t - N_s$  suit une loi de Poisson de paramètre  $\lambda(t - s)$ . Ce processus peut être obtenu en posant

$$N_t = \text{Card}\{k \geq 1 : T_k \leq t\} = \sum_{k \geq 1} 1_{T_k \leq t}$$

où  $T_k = T_{k-1} + E_k = E_1 + \dots + E_k$ , pour des v.a.  $(E_i)_{i \geq 1}$  i.i.d. de loi exponentielle de paramètre  $\lambda$ ; par convention,  $T_0 = 0$ .

Soit  $\lambda > 0$  et  $\nu$  une loi de probabilité sur  $\mathbb{R}$ . Un *processus de Poisson composé de paramètres  $(\lambda, \nu)$*  est un processus  $(X_t)_{t \geq 0}$  défini par

$$X_t = \sum_{k=1}^{N_t} Y_k$$

où les sauts  $(Y_k)_{k \geq 1}$  sont des v.a. réelles i.i.d. de loi  $\nu$  indépendantes de  $(N_t)_{t \geq 0}$ , processus de Poisson homogène de paramètre  $\lambda$ .

### 1.2. Simulations d'un processus de Poisson homogène.

**1.2.1. Simulation d'une trajectoire sur  $[0, T_n]$ ,  $n$  fixé.** Utiliser le fait que

$$N_t = \begin{cases} 0 & \text{pour } t \in [0, T_1[, \\ 1 & \text{pour } t \in [T_1, T_2[, \\ 2 & \text{pour } t \in [T_2, T_3[, \\ \vdots & \vdots \end{cases}$$

et que les instants de sauts  $(T_k)_{1 \leq k \leq n}$  s'expriment à l'aide de tirages de loi exponentielle. Pour l'affichage d'une fonction en escalier, on pourra utiliser la fonction `step` de `matplotlib.pyplot` (nous vous conseillons de consulter la documentation de cette fonction, notamment pour comprendre le rôle de son argument “where”).

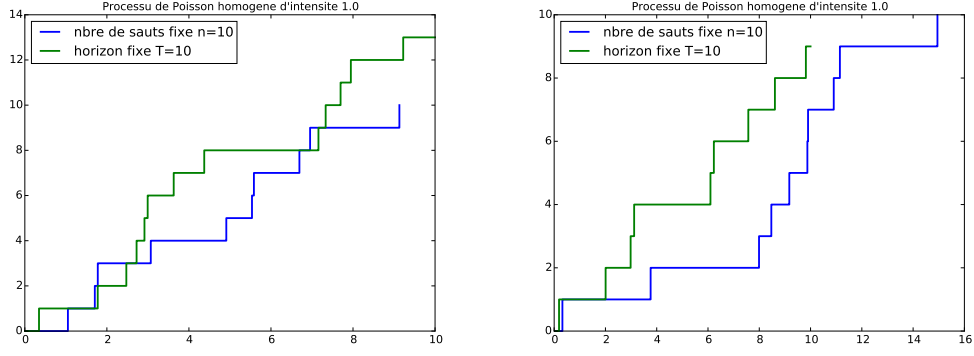


FIGURE 1. (question 1)  $\lambda = 1$ ,  $T = 10$  et  $n = 10$ .

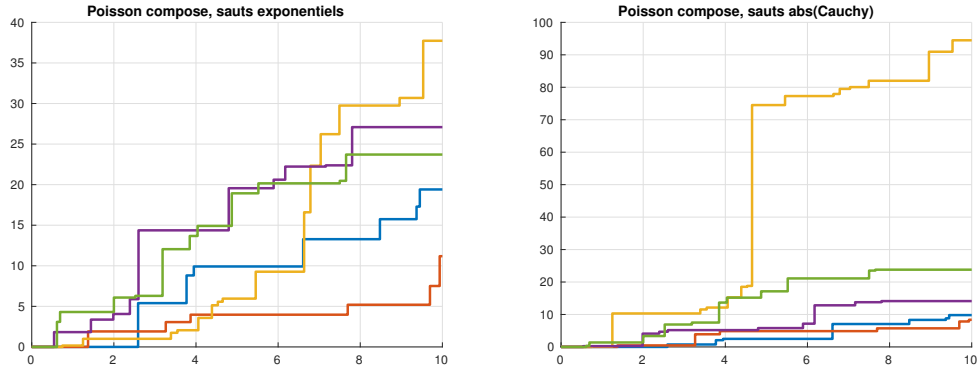


FIGURE 2. (question 2)  $\lambda = 1$  et  $T = 10$ . Cas exponentiel de moyenne 2 (gauche) et cas Cauchy standard (droite). Pour chacun des cas, on trace 5 trajectoires (mêmes instants de saut à gauche et à droite pour les courbes de même couleur).

1.2.2. *Simulation d'une trajectoire sur  $[0, T]$ ,  $T$  fixé.* Utiliser la loi de  $N_T$ ; puis le fait que conditionnellement à l'événement  $\{N_T = k\}$ , la loi des instants de saut  $(T_1, \dots, T_k)$  est la loi de  $k$  v.a. uniformes sur  $[0, T]$  réordonnées (voir Cours 1).

Q1) En utilisant ces deux approches, simuler et tracer une trajectoire d'un processus de Poisson homogène jusqu'à l'instant  $T_n$  (pour  $n \geq 1$  fixé) puis jusqu'à l'instant  $T$  (pour  $T > 0$  fixé). **Corrigé : Deux trajectoires pour chacun des cas, figure 1.**

Q2) Adapter le code précédent pour tracer la trajectoire d'un processus de Poisson composé; pour la loi des sauts  $(Y_k)_{k \geq 1}$ , on envisagera successivement une loi exponentielle, puis la loi de la valeur absolue d'une v.a. de Cauchy. Quelle différence remarque-t-on dans la régularité de l'amplitude des sauts ?

**Corrigé : Voir Figure 2. Trajectoires pour le cas exponentiel (gauche) de paramètre  $\mu = 1/2$ , et pour une loi de Cauchy standard (droite). Dans le cas d'un horizon fixe.**

Pour comparer les processus de Poisson composé à loi de sauts différente, les instants de sauts sont les mêmes sur les courbes de même couleur.

Cas Cauchy : la taille des sauts peut être très grande; la loi de Cauchy n'a pas de moments d'ordre 1.

## 2. CHANGEMENT DE PROBABILITÉ POUR LES PROCESSUS DE POISSON

Soient  $\lambda > 0$ ,  $\nu$  une loi de probabilité sur  $\mathbb{R}$ , des v.a. réelles  $(Y_k)_{k \geq 1}$  et un processus  $(N_t)_{t \geq 0}$  définis sur l'espace probabilisable  $(\Omega, \mathcal{F})$ . On munit cet espace de la probabilité  $\mathbb{P}$  telle que sous  $\mathbb{P}$ , les v.a.  $(Y_k)_{k \geq 1}$  sont i.i.d. de loi  $\nu$  et indépendantes de  $(N_t)_{t \geq 0}$ , processus de Poisson homogène de paramètre  $\lambda$ . Notons

$$X_t = \sum_{k=1}^{N_t} Y_k.$$

D'après la Section 1.1, sous  $\mathbb{P}$ ,  $(X_t)_{t \geq 0}$  est un processus de Poisson composé de paramètres  $(\lambda, \nu)$ . Dans la suite,  $\mathbb{E}$  et  $\text{Var}$  désignent resp. l'espérance et la variance sous  $\mathbb{P}$ .

**2.1. Rappel de cours : la transformation de Esscher.** À toute fonction mesurable  $f : \mathbb{R} \rightarrow \mathbb{R}$  telle que  $\mathbb{E}[\exp(f(Y_1))] < \infty$ , on associe une loi de probabilité  $\mathbb{P}_f$  sur  $(\Omega, \mathcal{F})$  définie par

$$(1) \quad \mathbb{P}_f(A) = \mathbb{E} \left[ 1_A \exp \left\{ \sum_{k=1}^{N_T} f(Y_k) - \lambda T (\mathbb{E}[\exp(f(Y_1))] - 1) \right\} \right],$$

pour tout  $A \in \mathcal{F}_T = \sigma(X_s, s \leq T)$ . Sous  $\mathbb{P}_f$ ,  $(X_t)_{t \geq 0}$  est encore un processus de Poisson composé, mais de paramètres  $(\lambda^f, \nu^f)$  donnés par

$$\lambda^f = \lambda \mathbb{E}[\exp(f(Y_1))], \quad \nu^f(dy) = \frac{\exp(f(y))}{\mathbb{E}[\exp(f(Y_1))]} \nu(dy).$$

**2.2. Cas de sauts constants.** Dans cette section, on s'intéresse au calcul de  $\mathbb{P}(X_T > x)$  pour des grandes valeurs  $x$  dans le cas où  $\nu$  est la mesure de Dirac en  $\alpha$ ,  $\alpha \in \mathbb{R}$  :  $\mathbb{P}(Y_1 = \alpha) = 1$ . On a donc  $X_t = \alpha N_t$  pour tout  $t \geq 0$ .

Dans toute la suite, pour  $\theta \in \mathbb{R}$ , on notera resp.  $\mathbb{P}_\theta, \nu^\theta, \lambda^\theta$  les quantités  $\mathbb{P}_f, \nu^f, \lambda^f$  introduites en Section 2.1 lorsque  $f : y \mapsto \theta y$ .  $\mathbb{E}_\theta$  est l'espérance sous  $\mathbb{P}_\theta$ .

Q3) Pour  $\theta \in \mathbb{R}$ , calculer la fonction log-laplace  $\Gamma_T(\theta)$  associée à  $X_T$ , définie par  $\Gamma_T(\theta) = \log \mathbb{E}[\exp(\theta X_T)]$ . En déduire l'espérance et la variance de  $X_T$  sous  $\mathbb{P}$ .

**Corrigé :**

$$\begin{aligned} \Gamma_T(\theta) &= \log \sum_{k \geq 0} \frac{(\lambda T)^k}{k!} \exp(-\lambda T) \exp(\alpha \theta k) = \lambda T (\exp(\alpha \theta) - 1), \\ \mathbb{E}[X_T] &= \Gamma'_T(0) = \lambda T \alpha, \\ \text{Var}(X_T) &= \Gamma''_T(0) = \lambda T \alpha^2. \end{aligned}$$

Q4) Exprimer le changement de loi (1) à l'aide de  $X_T, \theta$  et  $\Gamma_T(\theta)$ . Quelle est la loi de  $(X_t)_{t \geq 0}$  sous  $\mathbb{P}_\theta$ , pour  $\theta \in \mathbb{R}$  ?

**Corrigé :** Pour tout  $A \in \sigma(N_s, s \leq T)$ ,

$$\mathbb{P}_\theta(A) = \mathbb{E}[1_A \exp(\alpha \theta N_T - \Gamma_T(\theta))] = \mathbb{E}[1_A \exp(\theta X_T - \Gamma_T(\theta))]$$

et d'après le rappel, sous  $\mathbb{P}_\theta$ ,  $(X_t)_{t \geq 0}$  est un processus de Poisson composé de paramètres

$$\lambda^\theta = \lambda \exp(\alpha \theta), \quad \nu^\theta \equiv \nu = \delta_\alpha.$$

Avec ce changement de loi, on ne change pas la loi des sauts, mais on augmente (resp. diminue) le nombre moyen de sauts en choisissant le signe de  $\theta$  (précisément  $\theta > 0$ , resp.  $\theta < 0$ , si  $\alpha$  est positif).

Q5) Soit  $\theta \in \mathbb{R}, T > 0$ . En déduire l'expression d'une v.a.  $L_{\theta, T}$  telle que pour toute fonction mesurable positive  $g$  on a

$$(2) \quad \mathbb{E}[g(X_T)] = \mathbb{E}_\theta \left[ g(X_T) \frac{1}{L_{\theta, T}} \right].$$

**Corrigé :** De la question précédente, il vient

$$\mathbb{E}[g(X_T)] = \mathbb{E}_\theta [g(X_T) \exp(-\theta X_T + \Gamma_T(\theta))]$$

donc  $L_{\theta, T} = \exp(\theta X_T - \Gamma_T(\theta))$ .

Q6) Les valeurs de  $x$  qui rendront l'événement  $\{X_T > x\}$  rare sous  $\mathbb{P}$  seront de la forme

$$(3) \quad x = \mathbb{E}[X_T] + \rho \sqrt{\text{Var}(X_T)},$$

pour  $\rho > 0$  grand. Pour  $x > \mathbb{E}[X_T]$ , résoudre l'équation  $\Gamma'_T(\theta) = x$  sur  $\mathbb{R}$ . On notera  $\theta_x$  cette solution.

**Corrigé :** On trouve  $\theta_x = \alpha^{-1} \log(x/(\lambda T \alpha))$ .

Tous les calculs précédents montrent que, quitte à changer  $\lambda$  en  $\lambda T$ , on peut supposer que  $T = 1$ .

**2.3. Application.** Les processus de Poisson composés sont souvent utilisés pour modéliser l'état de la comptabilité des compagnies d'assurance. En effet, si une compagnie d'assurance de capital initial  $C$  reçoit à chaque unité de temps des cotisations d'une valeur de  $\Lambda$  et doit effectuer des dédommagements de loi  $\nu$  répartis dans le temps selon un processus de Poisson d'intensité  $\lambda$ , alors ses liquidités, à toute date  $t$ , sont données par

$$C + \Lambda t - X_t,$$

où  $X_t$  est un processus de Poisson composé d'intensité  $\lambda$  et de sauts de loi  $\nu$ . Afin d'éviter le risque de faillite, la compagnie d'assurance est alors amenée à évaluer  $\mathbb{P}(X_T > x)$  pour des valeurs de  $x$  élevées.

On prendra  $T = 1$ ,  $\lambda = 1$ ,  $\nu = \delta_\alpha$  pour  $\alpha = 0.5$  et  $x$  donné par (3) pour  $\rho = 1$ ,  $\rho = 5$  puis  $\rho = 20$ .

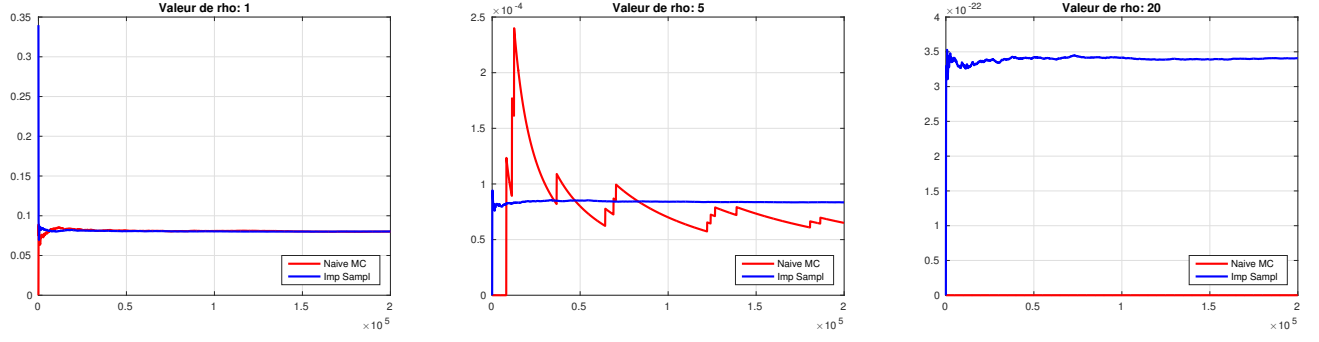


FIGURE 3. (question 7) Estimateur de Monte Carlo naïf (rouge) et d'échantillonnage d'importance (bleu) dans le cas  $\rho = 1$ ,  $\rho = 5$  et  $\rho = 20$ . On trace l'évolution de l'estimateur en fonction du nombre de points  $M$  dans l'échantillon.

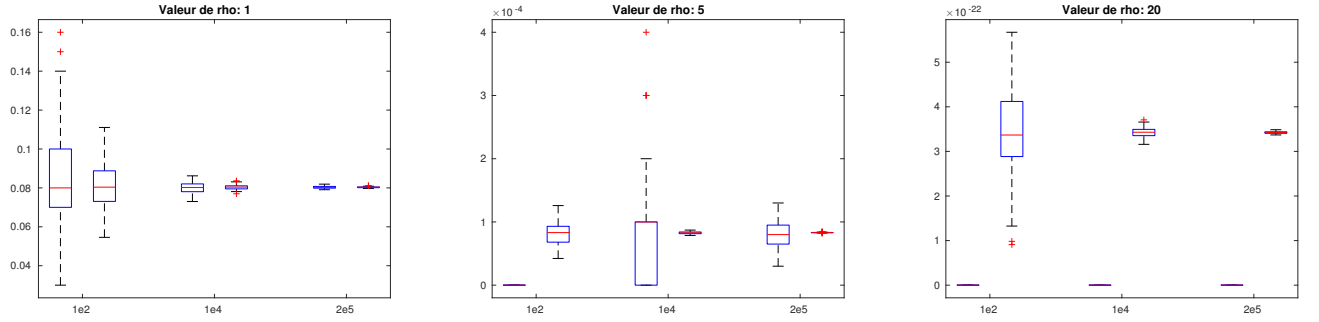


FIGURE 4. (question 7) Pour différentes valeurs de  $M$  ( $M = 1e2, 1e4, 2e5$ ), on trace le boxplot de 100 réalisations indépendantes des estimateurs  $\hat{\pi}_M^{MC}$  (boxplot gauche) et  $\hat{\pi}_M^{IS}$  (boxplot droit). Dans le cas  $\rho = 20$ , la méthode de Monte Carlo naïve échoue totalement et renvoie systématiquement la valeur zero comme estimateur.

Q7) Comparer la méthode de Monte Carlo naïve et la méthode d'échantillonnage d'importance issue de la relation (2) lorsque  $\theta = \theta_x$  : observer la consistance des estimateurs et donner un intervalle de confiance asymptotique de la quantité  $\mathbb{P}(X_T > x)$ .

**Corrigé :** On compare les estimateurs

$$\hat{\pi}_M^{MC} = \frac{1}{M} \sum_{k=1}^M 1_{N_T^{(k)} > x/\alpha}$$

$$\hat{\pi}_M^{IS} = \frac{1}{M} \sum_{k=1}^M 1_{\tilde{N}_T^{(k)} > x/\alpha} \exp(-\theta_x \alpha \tilde{N}_T^{(k)} + \Gamma_T(\theta_x))$$

avec  $\{N_T^{(k)}, k \geq 1\}$  v.a. i.i.d. Poisson de paramètre  $\lambda T$ ; et  $\{\tilde{N}_T^{(k)}, k \geq 1\}$  v.a. i.i.d. Poisson de paramètre  $\lambda T \exp(\alpha \theta_x)$ .  
Figure 3 : on trace l'évolution de la suite  $\{\hat{\pi}_M^\bullet, M \geq 1\}$  afin d'illustrer la consistance des estimateurs dans le cas  $\rho = 1$  (gauche),  $\rho = 5$  (centre) et  $\rho = 20$  (droite).

On calcule les intervalles de confiance asymptotique à 95% par la formule

$$\left[ \hat{\pi}_M^\bullet \pm 1.96 \frac{\sigma_M^\bullet}{\sqrt{M}} \right]$$

où

$$\left( \hat{\sigma}_M^{MC} \right)^2 = \hat{\pi}_M^{MC} \left( 1 - \hat{\pi}_M^{MC} \right)$$

$$\left( \hat{\sigma}_M^{IS} \right)^2 = \frac{1}{M} \sum_{k=1}^M 1_{\tilde{N}_T^{(k)} > x/\alpha} \exp(-2\theta_x \alpha \tilde{N}_T^{(k)} + 2\Gamma_T(\theta_x)) - \left( \hat{\pi}_M^{IS} \right)^2$$

Figure 4 : on calcule 100 réalisations indépendantes des estimateurs MC et IS. On affiche les boxplots pour différentes valeurs de  $M$  ( $M = 1e2, 1e4, 2e5$ ). On peut vérifier la plus petite variabilité de l'estimateur d'échantillonnage d'importance et sa plus grande vitesse de convergence; cette robustesse est notamment claire dans la situation  $\rho = 5$  où l'approche Monte Carlo commence à montrer quelques faiblesses bien qu'apportant encore une estimation non nulle pour des valeurs suffisamment grandes du nombre de points  $M$ . Dans le cas  $\rho = 20$ , la méthode de Monte Carlo échoue totalement en renvoyant 100 estimations égales à zero.

### 3. MÉTHODE DE SPLITTING ET PROCESSUS DE POISSON

**3.1. Rappels de cours : méthode de splitting via MCMC.** Soit  $\mathbf{N} = (N_t)_{t \geq 0}$  un processus de Poisson homogène de paramètre  $\lambda$ . Pour calculer la probabilité d'un événement rare de la forme  $\{\Phi(\mathbf{N}) \leq a\}$  où  $\Phi(\mathbf{N})$  désigne une fonctionnelle de la trajectoire du processus de Poisson  $(N_t)_{t \geq 0}$ , la méthode de splitting consiste à

- (i) choisir  $K$  seuils:  $a = a_K < \dots < a_2 < a_1 < a_0 = +\infty$ ;
- (ii) pour tout  $k = 1, \dots, K$ , approcher  $\mathbb{P}(\Phi(\mathbf{N}) \leq a_k | \Phi(\mathbf{N}) \leq a_{k-1})$  par un estimateur  $\hat{\pi}_k$ .
- (iii) estimer  $\mathbb{P}(\Phi(\mathbf{N}) \leq a)$  par  $\prod_{k=1}^K \hat{\pi}_k$ .

Soit  $k \in \{2, \dots, K\}$  fixé. L'algorithme *MCMC pour loi conditionnelle* ci-dessous permet de produire une suite de processus de Poisson  $\{\mathbf{N}^\ell, \ell \geq 0\}$ , vérifiant une loi forte des grands nombres: pour tout  $b \in \mathbb{R}$ , on a

$$\frac{1}{M} \sum_{\ell=1}^M \mathbf{1}_{\Phi(\mathbf{N}^\ell) \leq b} \xrightarrow[M \rightarrow \infty]{\text{p.s.}} \mathbb{P}(\Phi(\mathbf{N}) \leq b | \Phi(\mathbf{N}) \leq a_{k-1}).$$

Précisément, cet algorithme construit une chaîne de Markov  $\mathbf{N}^\ell$  à valeurs dans un espace de processus de Poisson, pour laquelle le théorème ergodique s'applique.<sup>1</sup> Il peut donc être utilisé pour construire l'estimateur  $\hat{\pi}_k$ .

**Algorithme MCMC pour loi conditionnelle  $\mathbb{P}(\cdot | \Phi(\mathbf{N}) \leq a_{k-1})$**

- Paramètres d'implémentation :  $p \in ]0, 1[$ ,  $a_{k-1} \in \mathbb{R}$ ,  $M \in \mathbb{N}$
- Initialisation :  $\mathbf{N}^0$  un processus de Poisson homogène de paramètre  $\lambda$  tel que  $\Phi(\mathbf{N}^0) \leq a_{k-1}$ .
- Pour  $\ell = 0, \dots, M-1$  : étant donné  $\mathbf{N}^\ell$ ,
  - construire un candidat  $\tilde{\mathbf{N}}^{\ell+1}$  par *coloriage*: (a) garder chaque saut de  $\mathbf{N}^\ell$  avec probabilité  $p$ , de manière indépendante: on obtient  $\mathbf{N}^{\ell,p}$ ; (b) construire un processus de Poisson homogène  $\tilde{\mathbf{N}}$  de paramètre  $(1-p)\lambda$ , indépendant de  $\mathbf{N}^{\ell,p}$ ; (c) renvoyer la somme  $\mathbf{N}^{\ell,p} + \tilde{\mathbf{N}}$ .
  - Si  $\Phi(\tilde{\mathbf{N}}^{\ell+1}) \leq a_{k-1}$ , poser  $\mathbf{N}^{\ell+1} = \tilde{\mathbf{N}}^{\ell+1}$ . Sinon, poser  $\mathbf{N}^{\ell+1} = \mathbf{N}^\ell$ .

**3.2. Application.** On reprend l'application décrite en Section 2.3, avec  $\nu = \delta_\alpha$ .

Q8) Montrer que p.s.

$$\min_{0 \leq t \leq T} \{C + \Lambda t - \alpha N_t\} = \min_{n \geq 0} \min_{t.q. T_n \leq T} \{C + \Lambda T_n - \alpha n\}.$$

**Corrigé :** En notant que pour tout  $t \in ]T_j, T_{j+1}[$ ,

$$C + \Lambda t - \alpha N_t = C + \Lambda t - \alpha j = (C + \Lambda T_j - \alpha j) + \Lambda(t - T_j) \geq C + \Lambda T_j - \alpha j$$

on en déduit l'égalité. Par conséquent, pour calculer le minimum des liquidités le long d'une trajectoire sur  $[0, T]$ , il suffit de regarder les valeurs prises par le processus des liquidités en les instants de sauts sur  $[0, T]$ .

Q9) Implémenter l'algorithme Monte Carlo naïf et l'algorithme de splitting combiné à l'algorithme *MCMC pour loi conditionnelle* pour approcher la probabilité que la ruine de l'assurance survienne avant l'instant  $T$ :

$$\mathbb{P} \left( \min_{0 \leq t \leq T} \{C + \Lambda t - \alpha N_t\} \leq 0 \right).$$

On prendra  $T = 1$ ,  $C = 100$ ,  $\Lambda = 30$ ,  $\alpha = 30$ ,  $a_k = C(1 - (k/K)^2)$  et successivement  $(\lambda, K) = (1, 3)$  et  $(\lambda, K) = (0.05, 8)$ . On envisagera différentes valeurs de  $p$ :  $p = 0.1$ ,  $p = 0.5$  et  $p = 0.9$ .

Pour le calcul de  $\hat{\pi}_k$ : dans le cas  $\lambda = 1$ , on pourra prendre  $M = 1e4$  quelle que soit la valeur de  $p$ ; et dans le cas  $\lambda = 0.05$ , on pourra prendre  $M = 8e4$  pour  $p = 0.5$ ; et  $M = 1.5e5$  pour  $p = 0.1$  et  $p = 0.9$ .

► Pour la méthode de splitting :

- a) Observer la consistance des estimateurs  $\hat{\pi}_k$ ,  $k = 1, \dots, K$  (c'est-à-dire, que ces sont des estimateurs convergents lorsque  $M$  devient grand).
- b) L'algorithme *MCMC pour loi conditionnelle* est un algorithme de rejet : on rejette tous les candidats  $\tilde{\mathbf{N}}^{\ell+1}$  qui ne satisfont pas à la condition  $\Phi(\mathbf{N}) \leq a_{k-1}$ . Pour chacune des  $K-1$  chaînes simulées par cet algorithme, il est intéressant d'observer l'évolution du taux d'acceptation le long de la trajectoire

$$M \mapsto \frac{1}{M} \sum_{\ell=1}^M \mathbf{1}_{\mathbf{N}^\ell = \tilde{\mathbf{N}}^\ell}.$$

Afficher l'évolution du taux d'acceptation.

<sup>1</sup>Pour plus de détails, on pourra consulter l'article E.Gobet et G.Liu (2015, SIAM Journal on Scientific Computing).

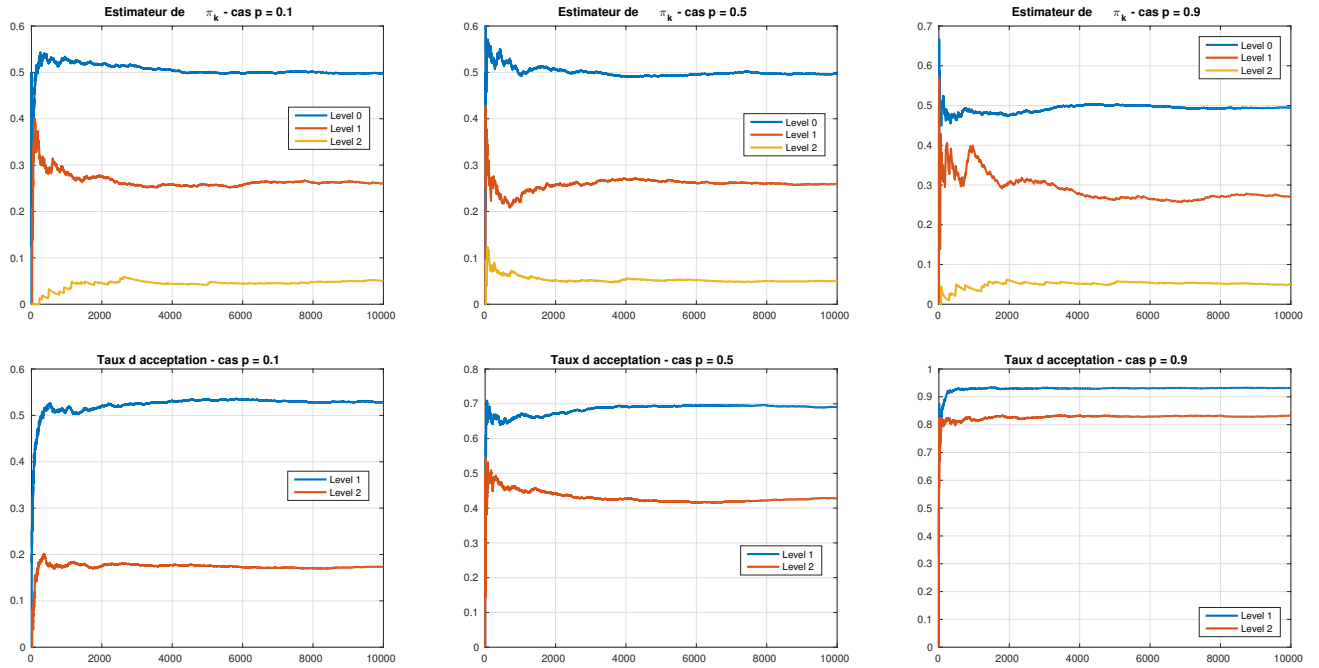


FIGURE 5. (question 9) Cas  $\lambda = 1$ . Consistance des estimateurs  $\hat{\pi}_k$  (haut) et évolution du taux d'acceptation le long de chaque chaîne (bas), en fonction de la longueur de la chaîne  $M$ . On considère différentes valeurs de  $p$  (gauche à droite).

- c) Estimer la variance de l'estimateur de la probabilité de ruine à l'aide de plusieurs réalisations indépendantes de cet estimateur. Calculer l'erreur relative i.e. le ratio de l'écart type et de la valeur moyenne de ces réalisations indépendantes.
- d) Quel est le meilleur choix de  $p$  ? Expliquer ce résultat.

► Comparer l'efficacité de l'estimateur Monte Carlo naïf et de l'estimateur par splitting pour le calcul de la probabilité de ruine.

**Corrigé :** Pour une comparaison plus juste, on compare l'estimateur de splitting avec  $K$  niveaux et  $M$  tirages par niveaux, à l'estimateur de Monte Carlo naïf avec  $KM$  tirages.

Dans le cas  $\lambda = 0.05$ , l'estimateur Monte Carlo naïf estime la probabilité de ruine par zero dans environ 94% des cas (calcul basé sur 900 réalisations indépendantes de cet estimateur). L'estimateur basé sur le splitting retourne une valeur strictement positive, de l'ordre de  $5.3 \cdot 10^{-8}$ . Sur la figure 7 (droite), on reporte le boxplot de 100 réalisations indépendantes de cet estimateur en fonction de  $p$ .

Sur les figures 5 et 6, on analyse une trajectoire de l'algorithme de splitting : (haut) on trace l'estimateur  $\hat{\pi}_k$  en fonction du nombre de points  $M$  de la chaîne; (bas) on trace l'évolution du taux d'acceptation le long de chaque chaîne. Lorsque  $p = 0.9$ , les taux d'acceptation sont très élevés; mais le point proposé  $\tilde{N}^{\ell+1}$  diffère peu du point courant  $N^\ell$ . Par suite, pour les  $K - 1$  chaînes : les valeurs successives d'une chaîne sont très corrélées; et dans le cas  $\lambda = 0.05$ , il est flagrant que cette forte corrélation ralentit la convergence de la moyenne empirique  $\hat{\pi}_k$ . Lorsque  $p = 0.1$ , le taux d'acceptation est assez faible, puisque le candidat  $\tilde{N}^{\ell+1}$  diffère beaucoup du point courant  $N^\ell$ : la chaîne reste au même point pendant plusieurs itérations. La corrélation entre valeurs successives est là aussi élevée. Il n'existe pas de résultats théoriques indiquant quel est le bon choix de  $p$ . Sur la figure 7(gauche), on trace l'évolution en fonction de  $p$ , de l'erreur relative (ratio de l'écart-type sur la moyenne) calculée à partir de 100 réalisations indépendantes de l'estimateur  $\prod_{k=1}^K \hat{\pi}_k$  de la probabilité de ruine. On peut vérifier que des valeurs trop faibles ou trop élevées de  $p$  ne sont pas optimales. Ici, l'erreur relative minimale est atteinte avec  $p = 0.6$  : l'écart type moyen est d'environ 0.37 fois la valeur moyenne de l'estimateur. Pour la méthode de Monte Carlo, on calcule 900 réalisations de l'estimateur et l'erreur relative estimée est de 4.1.

---

Pour aller plus loin

---

#### 4. CHANGEMENT DE PROBABILITÉ POUR LES PROCESSUS DE POISSON COMPOSÉS (SUITE)

4.1. **Cas de sauts de loi exponentielle.** Dans cette section, on s'intéresse au calcul de  $\mathbb{P}(X_T > a)$  pour des grandes valeurs  $a$  dans le cas où  $\nu$  est la loi exponentielle de paramètre  $\mu > 0$ :  $d\nu(y) = \mu \exp(-\mu y) \mathbf{1}_{\mathbb{R}^+}(y) dy$ .

Q10) Identifier  $\mathcal{D} = \{\theta \in \mathbb{R} : \int \exp(\theta y) d\nu(y) < \infty\}$ .

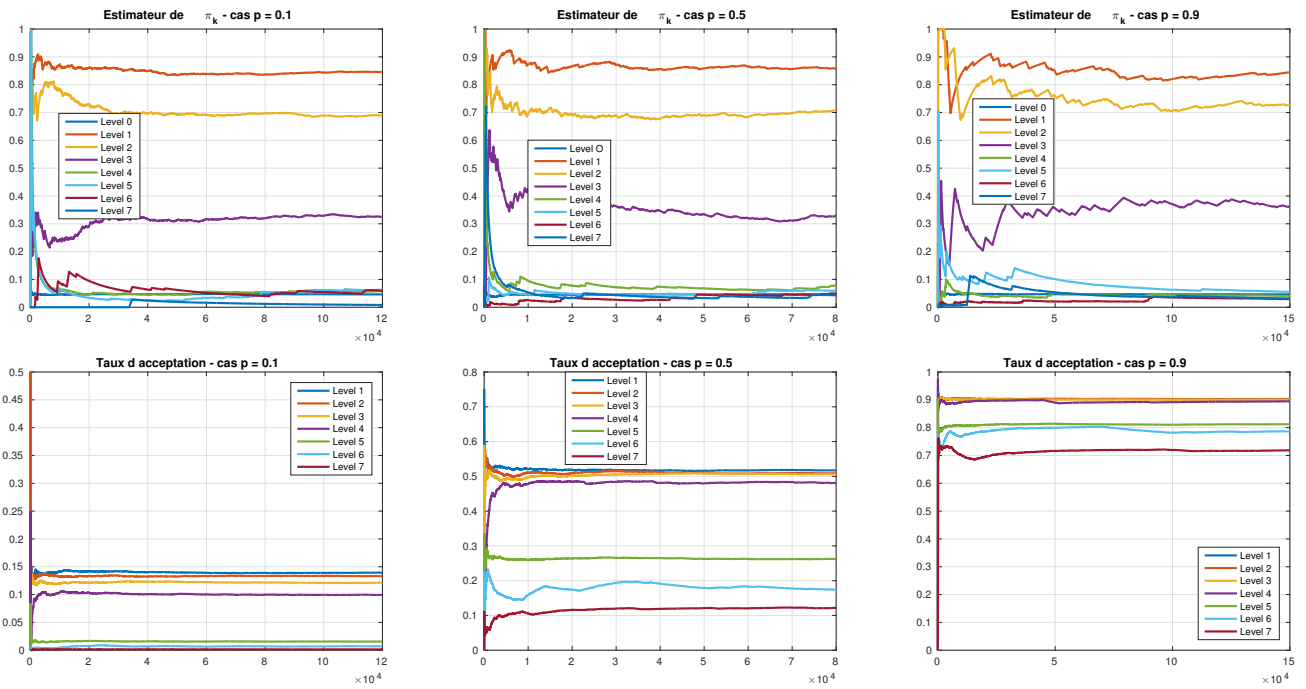


FIGURE 6. (question 9) Cas  $\lambda = 0.05$ . Consistance des estimateurs  $\hat{\pi}_k$  (haut) et évolution du taux d'acceptation le long de chaque chaîne (bas), en fonction de la longueur de la chaîne  $M$ . On considère différentes valeurs de  $p$  (gauche à droite).

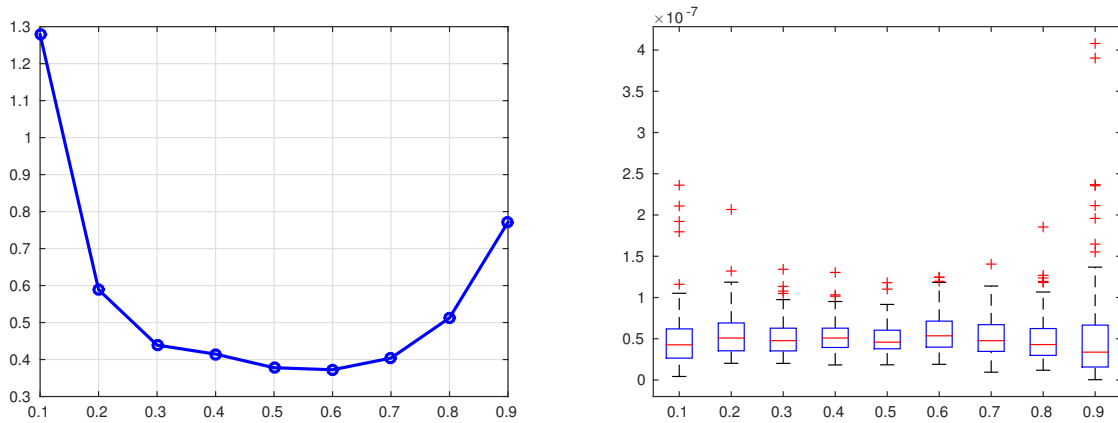


FIGURE 7. (question 9) Cas  $\lambda = 0.05$ ,  $K = 8$  et  $M = 1.5e5$ . (gauche) Evolution de l'erreur relative en fonction de  $p$ . (droite) Boxplot de 100 réalisations indépendantes de l'estimateur de splitting, pour différentes valeurs de  $p$ .

Corrigé : On a

$$\int \exp(\theta y) d\nu(y) = \frac{\mu}{\mu - \theta} \left( 1 - \lim_{y \rightarrow +\infty} \exp(-(\mu - \theta)y) \right)$$

dont on déduit que  $\mathcal{D} = ]-\infty, \mu[$ .

Q11) Pour  $\theta \in \mathcal{D}$ , calculer la fonction log-laplace  $\Gamma_T(\theta)$  associée à  $X_T$ , définie par  $\Gamma_T(\theta) = \log \mathbb{E}[\exp(\theta X_T)]$ . En déduire l'espérance et la variance de  $X_T$  sous  $\mathbb{P}$ .

Corrigé :

$$\Gamma_T(\theta) = \lambda T \left( \int \exp(\theta y) d\nu(y) - 1 \right) = \frac{\lambda T \theta}{\mu - \theta},$$

$$\mathbb{E}[X_T] = \Gamma'_T(0) = \frac{\lambda T}{\mu},$$

$$\text{Var}(X_T) = \Gamma''_T(0) = \frac{2\lambda T}{\mu^2}.$$

Q12) Exprimer le changement de loi (1) à l'aide de  $X_T, \theta$  et  $\Gamma_T(\theta)$ . Quelle est la loi de  $(X_t)_{t \geq 0}$  sous  $\mathbb{P}_\theta$ , pour  $\theta \in \mathcal{D}$  ?

**Corrigé :** Pour tout  $A \in \sigma(X_s, s \leq T)$ ,

$$\mathbb{P}_\theta(A) = \mathbb{E}[1_A \exp(\theta X_T - \Gamma_T(\theta))]$$

et d'après le rappel, sous  $\mathbb{P}_\theta$ ,  $(X_t)_{t \geq 0}$  est un processus de Poisson composé de paramètres

$$\lambda^\theta = \frac{\lambda \mu}{\mu - \theta}, \quad \nu^\theta \equiv \mathcal{E}(\mu - \theta).$$

Avec ce changement de loi, on augmente la taille moyenne des sauts ( $\mu - \theta < \mu$ ) et on augmente le nombre moyen de sauts ( $\lambda^\theta > \lambda$ ).

Q13) Les valeurs de  $a$  qui rendront l'événement  $\{X_T > a\}$  rare sous  $\mathbb{P}$  seront de la forme

$$a = \mathbb{E}[X_T] + \rho \sqrt{\text{Var}(X_T)},$$

pour  $\rho > 0$  grand. Pour  $a > \mathbb{E}[X_T]$ , résoudre l'équation  $\Gamma'_T(\theta) = a$  sur  $\mathcal{D}$ . On notera  $\theta_a$  cette solution.

**Corrigé :** On trouve  $\theta_a = \mu - \sqrt{\frac{\lambda T \mu}{a}}$

Tous les calculs précédents montrent que, quitte à changer  $\lambda$  en  $\lambda T$ , on peut supposer que  $T = 1$ .

4.2. **Application (suite).** Reprendre la question 7 dans le cas où  $\nu$  est la loi exponentielle de paramètre  $\mu$ . On prendra  $\mu = 1$ .