
Feuille de TP n°11

Loi exponentielle en modélisation

1 Géométrie, Exponentielle, Weibull

La loi géométrique \mathcal{G} , la loi exponentielle \mathcal{E} et la loi de Weibull \mathcal{W} sont souvent utilisées pour modéliser des temps d'attente ou des durées de vie. Il se trouve que \mathcal{G} est la version discrète de \mathcal{E} , et que \mathcal{E} est un cas particulier de \mathcal{W} .

Remarque 1.1. Dans toute la suite, la loi géométrique de paramètre $p \in]0, 1[$, notée $\mathcal{G}(p)$, désignera la loi $p \sum_{n=1}^{\infty} (1-p)^{n-1} \delta_n$ ou bien la loi $p \sum_{n=0}^{\infty} (1-p)^n \delta_n$, le passage de l'une à l'autre se faisant en translatant le support de 1. La définition utilisée sera précisée si nécessaire pour éviter les confusions.

Définition 1.2. Soit X une v.a.r. de fonction de répartition F sur \mathbb{R} . Sa loi est dite

1. exponentielle $\mathcal{E}(\lambda)$ avec $\lambda > 0$ si $F(t) = (1 - \exp(-\lambda t)) \mathbb{I}_{\{t \geq 0\}}$;
2. de Weibull $\mathcal{W}(a, \lambda)$ avec $a > 0$ et $\lambda > 0$ si $F(t) = (1 - \exp(-\lambda t^a)) \mathbb{I}_{\{t \geq 0\}}$.

Ces deux lois sont absolument continues, de densité respectives

$$x \mapsto \lambda \exp(-\lambda x) \mathbb{I}_{\{x \geq 0\}} \quad \text{et} \quad x \mapsto a \lambda x^{a-1} \exp(-\lambda x^a) \mathbb{I}_{\{x \geq 0\}}.$$

Si $X \sim \mathcal{W}(a, \lambda)$, alors $X^a \sim \mathcal{E}(\lambda)$. Si $a = 1$, alors X suit la loi exponentielle $\mathcal{E}(\lambda)$. Si $a = 2$, alors X suit la loi de Rayleigh $\mathcal{R}(a)$.

Remarque 1.3 (Premiers moments...). Les deux premiers moments de $\mathcal{E}(\lambda)$ sont λ^{-1} et λ^{-2} . Plus généralement, les deux premiers moments de $\mathcal{W}(a, \lambda)$ sont donnés par $\lambda^{-1/a} \Gamma(1+1/a)$ et $\lambda^{-2/a} (\Gamma(1+2/a) - \Gamma^2(1+1/a))$ où Γ est la fonction Gamma d'Euler définie par $\Gamma(a) := \int_0^{\infty} x^{a-1} \exp(-x) dx$.

Remarque 1.4 (Amnésie exponentielle). La loi exponentielle hérite de sa fonction de répartition exponentielle des propriétés remarquables. Ainsi, une v.a.r. X absolument continue à valeurs dans \mathbb{R}_+ suit une loi exponentielle si et seulement si, pour tout $s, t \geq 0$,

$$\mathbb{P}(X > t + s \mid X \geq t) = \mathbb{P}(X > s). \tag{1}$$

Dans le même esprit, la v.a.r. absolument continue X à valeur dans \mathbb{R}_+ suit une loi exponentielle si et seulement si, pour tout $s, t > 0$, $\mathbb{P}(X \leq t + s \mid X > t) = st + o(s)$. La propriété (1) exprime le fait que la loi exponentielle se caractérise par une *absence de mémoire* :

$$\forall t > 0, \mathcal{L}(X - t \mid X \geq t) = \mathcal{L}(X).$$

Cela traduit bien le comportement de durées de vie en fiabilité, du temps d'arrivée du prochain client dans une file d'attente, etc. Si X_1, \dots, X_n sont indépendantes et de lois respectives $\mathcal{E}(\lambda_1), \dots, \mathcal{E}(\lambda_n)$, alors $\min(X_1, \dots, X_n) \sim \mathcal{E}(\lambda_1 + \dots + \lambda_n)$ et $\mathbb{P}(X_i = \min(X_1, \dots, X_n)) = \lambda_i / (\lambda_1 + \dots + \lambda_n)$. Comme nous allons le voir, cette dernière propriété joue un rôle très important dans l'étude des files d'attentes ainsi que pour leur simulation.

Remarque 1.5 (Amnésie géométrique). Ici, on considère la loi géométrique $\mathcal{G}(p)$ de paramètre $p \in]0, 1[$ donnée par $\mathcal{G}(p) := p \sum_{n=0}^{\infty} (1-p)^n \delta_n$, mais ce qui est dit d'adapte sans immédiatement à l'autre définition. La fonction de répartition de $\mathcal{G}(p)$ est nulle sur $]-\infty, 0[$, tandis que sur les intervalles de la forme $[n-1, n[$ avec $n \in \mathbb{N}^*$, elle vaut $1 - (1-p)^n = 1 - e^{-n \log(1-p)}$. La loi $\mathcal{G}(p)$ hérite de la sa fonction de répartition quasi-exponentielle

des propriétés remarquables. Ainsi, une v.a.r. discrète X à valeurs dans \mathbb{N} suit une loi $\mathcal{G}(p)$ si et seulement si, pour tout $n, m \in \mathbb{N}$,

$$\mathbb{P}(X > n + m \mid X \geq n) = \mathbb{P}(X > m). \quad (2)$$

La propriété (2) exprime le fait que la loi géométrique se caractérise par une *absence de mémoire* :

$$\forall n \in \mathbb{N}^*, \mathcal{L}(X - n \mid X \geq n) = \mathcal{L}(X).$$

Cela traduit bien le comportement de durées de vie en fiabilité, du temps d'arrivée du prochain client dans une file d'attente, etc. Si X_1, \dots, X_n sont indépendantes et de lois respectives $\mathcal{G}(p_1), \dots, \mathcal{G}(p_n)$, alors $\min(X_1, \dots, X_n) \sim \mathcal{G}(1 - (1 - p_1) \cdots (1 - p_n))$.

Exercice 1.6 (Petit calcul). Soit X_1, \dots, X_n des v.a. indépendantes et de lois respectives $\mathcal{G}(p_1), \dots, \mathcal{G}(p_n)$ où $\mathcal{G}(p) := p \sum_{n=0}^{\infty} (1-p)^n \delta_n$. Exprimer $\mathbb{P}(X_i = \min(X_1, \dots, X_n))$ en fonction de p_1, \dots, p_n . Ce calcul sera utile à résolution de l'exercice 3.5.

Remarque 1.7 (Quasi-amnésie de Weibull). Si $X \sim \mathcal{W}(a, \lambda)$, alors pour tout $s, t \geq 0$,

$$\mathbb{P}(X > t + s \mid X > t) \begin{cases} \geq \mathbb{P}(X > s) & \text{si } a \leq 1, \\ \leq \mathbb{P}(X > s) & \text{si } a \geq 1. \end{cases}$$

Remarque 1.8 (La loi géométrique comme partie entière de la loi exponentielle). Pour tout $x \in \mathbb{R}_+$, on note $\lfloor x \rfloor$ sa partie entière et $\lceil x \rceil := x - \lfloor x \rfloor$ sa partie fractionnaire. Si $X \sim \mathcal{E}(\lambda)$, alors $\lfloor X \rfloor$ et $\lceil X \rceil$ sont indépendantes. De plus, $\lfloor X \rfloor \sim \mathcal{G}(1 - e^{-\lambda})$ où $\mathcal{G}(p) := p \sum_{n=0}^{\infty} (1-p)^n \delta_n$, et enfin $\lceil X \rceil$ suit la loi $\mathcal{L}(X \mid X \in [0, 1])$ de densité suivante:

$$x \in \mathbb{R} \rightarrow \frac{1}{1 - e^{-\lambda}} \lambda e^{-\lambda x} \mathbf{I}_{[0,1]}(x).$$

Ce fait¹ est une conséquence directe de l'absence de mémoire de $\mathcal{E}(\lambda)$, qui entraîne que $\mathcal{L}(\lceil X \rceil \mid \lfloor X \rfloor)$ ne dépend pas de $\lfloor X \rfloor$ et vaut $\mathcal{L}(X \mid X \in [0, 1])$. L'intuition à retenir est que dans un jeu de pile ou face de paramètre de gain $p \in]0, 1[$, la quantité $-\log(1-p)$ joue le rôle du paramètre λ d'une loi exponentielle. La loi géométrique est l'analogue discret de la loi exponentielle. Nous allons voir que la loi exponentielle apparaît comme la loi des temps inter-sauts d'un processus de Poisson, tandis que la loi géométrique apparaît comme la loi des temps inter-sauts d'un processus de Bernoulli.

Remarque 1.9 (Lois de Dirichlet – Statistiques d'ordre – Lois exponentielles). Soit (U_1, \dots, U_n) un vecteur de v.a.r. i.i.d. de loi uniforme sur $[a, b]$. Soit $V := (U_{(1)}, \dots, U_{(n)})$ son réarrangement croissant. La loi de V est appelée *loi de Dirichlet* $\mathcal{D}_n([a, b])$, ou *statistique d'ordre*. Elle est absolument continue par rapport à la mesure de Lebesgue sur \mathbb{R}^n et admet pour densité

$$\frac{n!}{(b-a)^n} \mathbf{I}_{\{a < x_1 < \dots < x_n < b\}}.$$

Soit à présent $(D_n)_{n \in \mathbb{N}^*}$ une suite de v.a.r. i.i.d. de loi exponentielle $\mathcal{E}(\lambda)$. Pour tout $n \in \mathbb{N}^*$, soit $T_n := D_1 + \dots + D_n$, qui suit une loi Gamma² $\Gamma(n, \lambda)$. On a alors pour tout $t > 0$,

$$\mathcal{L}(T_1, \dots, T_{n-1} \mid T_{n-1} < t < T_n) = \mathcal{L}(T_1, \dots, T_{n-1} \mid T_n = t) = \mathcal{D}_{n-1}([0, t]).$$

2 Processus de Poisson simple et processus de Bernoulli

Définition 2.1 (Processus de Poisson simple). On dit que $(N_t)_{t \geq 0}$ est un processus de Poisson simple d'intensité $\lambda > 0$ issu de zéro si et seulement si $N_0 = 0$ et pour tout $s, t > 0$, les v.a.r. $N_{t+s} - N_s$ et N_s sont indépendantes

¹Plus généralement, si $Y_t := \lfloor X/t \rfloor + 1$ avec $t > 0$ fixé et $X \sim \mathcal{E}(\lambda)$, alors $Y_t \sim \mathcal{G}(1 - \exp(-\lambda t))$, etc.

²De densité $\frac{1}{(n-1)!} t^{n-1} \exp(-\lambda t) dt$.

et de loi de Poisson $\mathcal{P}(\lambda s)$ et $\mathcal{P}(\lambda t)$ respectivement. Ainsi, le processus de Poisson simple est un processus à temps continu et à accroissements stationnaires et indépendants, tout comme le mouvement brownien³

Définition 2.2 (Processus de Bernoulli). On dit que $(B_n)_{n \in \mathbb{N}}$ est un processus de Bernoulli de paramètre $p \in]0, 1[$ issu de zéro si et seulement si $B_0 = 0$ et pour tout $n, m \in \mathbb{N}^*$, $B_{n+m} - B_m$ et B_m sont indépendantes et de loi binomiale $\mathcal{B}(m, p)$ et $\mathcal{B}(n, p)$ respectivement. Ainsi, le processus de Bernoulli est un processus à temps discret et à accroissements stationnaires et indépendants.

Remarque 2.3 (Simuler la trajectoire d'un processus de Poisson simple). Simuler une trajectoire à n sauts d'un processus de Poisson simple d'intensité λ est une tâche très simple:

```
stairs(cumsum(rexpweib(n, lambda, 1)), [1:n]);
```

Proposition 2.4 (Temps inter-sauts d'un processus de Poisson simple). Soit $(D_n)_{n \in \mathbb{N}^*}$ une suite de v.a.r. i.i.d. de loi exponentielle $\mathcal{E}(\lambda)$ avec $\lambda > 0$. Pour tout $t \geq 0$, si

$$N_t := \sum_{n=1}^{\infty} \mathbf{I}_{\{T_n \leq t\}} \quad \text{avec} \quad T_n := \sum_{k=1}^n D_k.$$

Alors $(N_t)_{t \geq 0}$ est un processus de Poisson simple issu de zéro, d'intensité λ . Réciproquement, les temps inter-sauts $(D_n)_{n \in \mathbb{N}^*}$ d'un processus de Poisson simple $(N_t)_{t \geq 0}$ d'intensité $\lambda > 0$ sont i.i.d. et suivent des lois exponentielles $\mathcal{E}(\lambda)$; quant aux temps de saut $(T_n)_{n \in \mathbb{N}^*}$, ils suivent des lois Gamma: le n -ième saut suit la loi Gamma $\mathcal{E}(\lambda)^{*n} = \Gamma(n, \lambda)$.

Proposition 2.5 (Temps inter-sauts d'un processus de Bernoulli). Soit $(D_n)_{n \in \mathbb{N}^*}$ une suite de v.a.r. i.i.d. de loi géométrique $\mathcal{G}(p) := p \sum_{n=1}^{\infty} (1-p)^{n-1} \delta_n$ avec $p \in]0, 1[$. Pour tout $n \in \mathbb{N}$, si

$$B_n := \sum_{m=1}^{\infty} \mathbf{I}_{\{T_m \leq n\}} \quad \text{avec} \quad T_m := \sum_{k=1}^m D_k.$$

Alors $(B_n)_{n \in \mathbb{N}}$ est un processus de Bernoulli issu de zéro, de paramètre p . Réciproquement, les temps inter-sauts $(D_n)_{n \in \mathbb{N}^*}$ d'un processus de Bernoulli $(B_n)_{n \in \mathbb{N}}$ de paramètre $p \in]0, 1[$ sont i.i.d. et suivent des lois géométriques $\mathcal{G}(p) = p \sum_{n=1}^{\infty} (1-p)^{n-1} \delta_n$; quant aux temps de saut $(T_n)_{n \in \mathbb{N}^*}$, ils suivent des lois négatives-binomiales⁴: le n -ième saut suit la loi $\mathcal{G}(p)^{*n}$.

Remarque 2.6 (Le jeu de pile ou face et le processus de Bernoulli). Si $(Z_n)_{n \in \mathbb{N}^*}$ est une suite de v.a. de loi de Bernoulli $\mathcal{B}(p) = (1-p)\delta_0 + p\delta_1$, alors le processus $(B_n)_{n \in \mathbb{N}}$ défini par $B_0 = 0$ et $B_n := Z_1 + \dots + Z_n$ est un processus de Bernoulli de paramètre p . Ainsi, B_n est le nombre de gains après le n -ième lancé à un jeu de pile ou face avec probabilité p de gagner.

Théorème 2.7 (Loi forte des grands nombres). L'indépendance et la stationnarité des accroissements des processus de Poisson et de Bernoulli entraîne une loi forte des grands nombres. Soit $(N_t)_{t \geq 0}$ un processus de Poisson simple d'intensité $\lambda > 0$, alors

$$\frac{N_t}{t} \xrightarrow[t \rightarrow +\infty]{p.s.} \lambda.$$

³Ils appartiennent tous deux à la classe des *processus de Lévy*, qui regroupe les processus à temps continu à accroissements stationnaires et indépendants. Un processus $(X_t)_{t \geq 0}$ appartenant à cette classe est caractérisé par la donnée de $\mathcal{L}(X_1)$, et $\mathcal{L}(X_t)$ est obtenue à partir de $\mathcal{L}(X_1)$ par la dilatation de coefficient t . Pour un processus de Lévy, $\mathcal{L}(X_1)$ est toujours *infiniment divisible*, i.e. pour tout $n \in \mathbb{N}^*$, il existe une loi ν_n telle que $\nu_n^{*n} = \mathcal{L}(X_1)$. Réciproquement, une loi infiniment divisible donne naissance à un processus de Lévy. Les lois infiniment divisibles sont caractérisées par une transformée de Fourier de forme très particulière (formule de Lévy-Khinchine). Les lois gaussiennes, les lois Gamma, et la loi de Poisson sont infiniment divisibles. Bien entendu, la notion d'infinie divisibilité n'a pas de sens pour les lois discrètes à support fini, mais d'une certaine façon, les lois binomiales et négatives-binomiales sont « finiment divisibles ».

⁴Appelées également lois de Pascal.

Soit $(B_n)_{n \in \mathbb{N}}$ un processus de Bernoulli de paramètre $p \in]0, 1[$, alors

$$\frac{B_n}{n} \xrightarrow[n \rightarrow +\infty]{p.s.} p.$$

Définition 2.8 (Processus de renouvellement et de comptage). Soit $(D_n)_{n \in \mathbb{N}^*}$ des v.a.r. i.i.d. strictement positives de loi μ . Le processus $(T_n)_{n \in \mathbb{N}}$ défini par $T_0 = 0$ et $T_n := D_1 + \dots + D_n$ est appelé processus de renouvellement de la suite $(D_n)_{n \in \mathbb{N}^*}$. Lorsque la loi μ est portée par \mathbb{N} , les v.a. $(D_n)_{n \in \mathbb{N}^*}$ sont à valeurs discrètes et on appelle processus de comptage associé à la suite $(T_n)_{n \in \mathbb{N}^*}$ le processus à temps discret $(N_n)_{n \in \mathbb{N}}$ défini par

$$N_n := \sum_{k=1}^{\infty} \mathbf{1}_{\{T_k \leq n\}}.$$

Lorsque μ est plus générale, on appelle processus de comptage associé à la suite $(T_n)_{n \in \mathbb{N}^*}$ le processus à temps continu $(N_t)_{t \geq 0}$ défini par

$$N_t := \sum_{k=1}^{\infty} \mathbf{1}_{\{T_k \leq t\}}.$$

Remarque 2.9. Le processus de Poisson simple d'intensité λ est le processus de comptage (à temps continu) associé à la loi exponentielle $\mathcal{E}(\lambda)$. Le processus de Bernoulli de paramètre p est le processus de comptage (à temps discret) associé à la loi géométrique $\mathcal{G}(p) := p \sum_{n=1}^{\infty} (1-p)^{n-1} \delta_n$.

3 Utilisation en modélisation

Exercice 3.1 (Records). Soit $(D_n)_{n \in \mathbb{N}^*}$ une suite de v.a.r. i.i.d. de loi exponentielle $\mathcal{E}(\lambda)$ avec $\lambda > 0$ et soit $M_n := \max_{1 \leq k \leq n} D_k$. Montrer que

$$\frac{M_n}{\log n} \xrightarrow[n \rightarrow +\infty]{p.s.} \frac{1}{\lambda} \quad \text{et} \quad M_n - \frac{\log n}{\lambda} \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{E}_{\mathcal{D}}(\lambda),$$

où $\mathcal{E}_{\mathcal{D}}(\lambda)$ est la loi exponentielle double de paramètre λ . Créer un code Matlab permettant d'illustrer ces résultats de convergence. La v.a.r. M_n correspond au plus long temps entre deux sauts avant le n -ième saut d'un processus de Poisson simple d'intensité λ .

Exercice 3.2 (Comptage). Créer un code Matlab permettant de simuler un processus de comptage $(N_t)_{t \geq 0}$ associé à la loi de Weibull $\mathcal{W}(a, \lambda)$ où les paramètres $a > 0$ et $\lambda > 0$ sont affectés par l'utilisateur. Vérifier la loi des grands nombres et le théorème de limite centrale pour le processus $(N_t)_{t \geq 0}$.

Exercice 3.3 (Renouvellement). Soit $(N_t)_{t \geq 0}$ un processus de Poisson simple d'intensité $\lambda > 0$ et soit $(T_n)_{n \in \mathbb{N}^*}$ un processus de renouvellement associé à une suite de v.a.r. i.i.d. strictement positives d'espérance m et de variance $\sigma^2 > 0$. On suppose que $(T_n)_{n \in \mathbb{N}^*}$ est indépendante de $(N_t)_{t \geq 0}$ et, pour tout $n \in \mathbb{N}$, on pose $Z_n := N_{T_n}$. Créer un code Matlab permettant de simuler les processus $(N_t)_{t \geq 0}$ et $(Z_n)_{n \in \mathbb{N}}$ où les paramètres λ , m et σ^2 sont affectés par l'utilisateur. Vérifier la loi des grands nombres et le théorème de limite centrale pour le processus $(Z_n)_{n \in \mathbb{N}}$.

Exercice 3.4 (Assureur communiste et réservoir percé). On peut modéliser le solde du compte en banque d'un assureur communiste à l'instant t par $x + \mu t - N_t$ où $\mu > 0$ et où $(N_t)_{t \geq 0}$ est un processus de Poisson simple d'intensité λ . La « dérive » linéaire μt représente les apports réguliers de ses assurés (cotisations) tandis que le processus de Poisson simple $(N_t)_{t \geq 0}$ modélise les catastrophes qui provoquent des remboursements. Pourquoi une telle modélisation n'est pas totalement farfelue ? Un assureur capitaliste préférera faire fructifier une partie de son solde en bourse, ce qui rend le modèle plus complexe. Comment tenir compte de la fluctuation du nombre d'assurés ? Comment adapter le modèle au cas où l'assureur place une partie de son solde en caisse d'épargne à taux fixe $\rho > 0$? Les lecteurs communistes ou rétifs aux assurances peuvent penser par exemple au compte en

banque d'une association. Exprimer le temps de ruine de l'assureur communiste en fonction des paramètres λ et μ , et confronter les prédictions à des simulations. Le processus $(\mu t - N_t)_{t \geq 0}$ permet également de modéliser la quantité d'eau dans un réservoir à ciel ouvert percé d'un trou, en plein hiver, en négligeant l'évaporation. Expliquer pourquoi.

Exercice 3.5 (Files d'attentes). Des clients attendent en file devant un guichet. On suppose que les temps séparant les arrivées de deux usagers successifs sont modélisables par des v.a.r. i.i.d. $\mathcal{E}(\lambda)$, tandis que les temps mis par le guichetier pour servir les clients successifs sont modélisables par des v.a.r. i.i.d. de loi $\mathcal{E}(\mu)$, indépendantes des précédentes. On parle de file d'attente M/M/1. Écrire une fonction Matlab simulant le fonctionnement d'un tel guichet. Vérifiez que lorsque $\rho := \lambda/\mu < 1$, le nombre d'usagers N_t dans la file d'attente converge en loi, lorsque le temps t tend vers l'infini, vers la loi géométrique $\mathcal{G}(\rho)$. Que représente la quantité ρ^K pour un entier K fixé (penser à une salle d'attente de K places : file d'attente M/M/1/K). Le paramètre ρ représente la *charge de la file*. Vérifiez que lorsque $\rho > 1$, N_t diverge presque sûrement. Étendre ce qui précède au cas où $s \in \mathbb{N}^*$ guichets indépendants sont disponibles (file M/M/s). On pourra également penser à des serveurs informatiques (resp. pistes d'aéroport) qui jouent le rôle de guichets, et à des programmes clients (resp. avions) qui jouent le rôle d'usagers. Après avoir pensé au problème, vérifiez que les deux codes Matlab qui suivent sont valides...

```
% ### Copyright (C) D. Chafai, 2001. GNU GPL.
% ### http://www.lsp.ups-tlse.fr/Chafai/agregation.html
% File d'attente M/M/s. Les clients arrivent avec des temps exp(1). Les s
% serveurs servent avec des temps exp(m). N(t) = nombre de clients dans la
% file au temps t y compris ceux qui sont en train d'être servis.
% Tester T=+Inf.
clear; clf;
s=2; % nombre de serveurs
l=.5; % intensité d'arrivée des clients
m=.26; % intensité du traitement d'un serveur
T=20; % on cherche la longueur de la file à cet instant
r=1000; % nombre de réalisations i.i.d.
for i=1:r,
    n=0; % la file est vide au depart
    t=-log(rand)/l; % on attend le premier client, temps de loi exp(1)
    while (t<T) % on boucle tant que le temps est < t
        % il s'est passé quelque chose
        if n==0 % la file était vide
            n=1; % un client de plus
        else % la file n'était pas vide
            % un client de plus avec proba l/(l+m*min(n,s))
            % un client de moins avec proba m*min(n,s)/(l+m*min(n,s))
            n=n+2*(rand<l/(l+m*min(n,s)))-1;
        end
        % on attend le prochain changement, temps de loi exp(l+m*min(n,s))
        t=t-log(rand)/(l+m*min(n,s));
    end
    N(i)=n;
end
histo(N)
[mean(N),std(N)]
```

Exercice 3.6 (Processus de Poisson composé et arbres de Galton-Watson). FIXME: à faire.

Exercice 3.7 (La machine est en panne). On considère une machine qui peut tomber en panne et qui est réparée instantanément. On peut, par exemple, penser à une machine avec une pile, une lampe avec une ampoule, etc.

Cette machine est mise en fonctionnement à l'instant 0. Soit D_1 sa durée de vie avant la première panne. Lorsqu'elle tombe en panne, on la remplace par une machine neuve identique de durée de vie D_2 et ainsi de suite. On suppose que $(D_n)_{n \in \mathbb{N}}$ est une suite de v.a.r i.i.d. strictement positives d'espérance m et de variance σ^2 . Pour tout $n \in \mathbb{N}$, T_n correspond à l'instant de la n -ième panne, avec $S_0 = 0$. Finalement, pour tout $t \geq 0$, N_t est le nombre de pannes survenues avant l'instant t et $H(t) := \mathbb{E}(N_t)$ est la fonction de renouvellement du processus $(N_t)_{t \geq 0}$. Montrer que si $(N_t)_{t \geq 0}$ est un processus de Poisson simple d'intensité $\lambda > 0$, $H(t) = \lambda t$. Plus généralement, montrer que $t/m - 1 \leq H(t) \leq t/m + \sigma^2/m^2$. Ensuite, vérifier que

$$\lim_{t \rightarrow \infty} \frac{H(t)}{t} = \frac{1}{m} \quad \text{et} \quad \lim_{t \rightarrow \infty} H(t) - \frac{t}{m} = \frac{\sigma^2 - m^2}{2m^2}.$$

Créer un code Matlab permettant d'illustrer les résultats de convergence associés à la fonction de renouvellement H , où les paramètres λ , m et σ^2 sont affectés par l'utilisateur.

Exercice 3.8 (Réseaux de Jackson et de Petri). FIXME: à faire. Cf. livre d'Ycart.

Exercice 3.9 (Processus ponctuel de Poisson et vie du charençon). FIXME: à faire. Cf. section 7.5 du livre de J. Istas.

4 Quelques extensions hors programme

Exercice 4.1 (Le générateur de Markov du processus de Poisson simple). Dans tout ce qui suit, λ est un réel positif fixé une fois pour toute. Soit $\mathcal{C}_b(\mathbb{R}, \mathbb{R})$ (resp. $\mathcal{C}_c(\mathbb{R}, \mathbb{R})$) l'espace vectoriel des fonctions réelles à valeurs réelles continues et bornées (resp. continues et à support compact). On considère l'application linéaire $\mathbf{L} : \mathcal{C}_b(\mathbb{R}, \mathbb{R}) \rightarrow \mathcal{C}_b(\mathbb{R}, \mathbb{R})$ définie pour tout $f \in \mathcal{C}_b(\mathbb{R}, \mathbb{R})$ et tout $x \in \mathbb{R}$ par

$$\mathbf{L}(f)(x) := \lambda(f(x+1) - f(x)).$$

Montrer que son noyau est constitué des fonctions constantes. On rappelle que $\mathbb{R}^{\mathbb{N}}$ est l'espace vectoriel des suites de réels, qui sont des fonctions de \mathbb{N} dans \mathbb{R} , et donc des vecteurs infinis. Pour tout $x \in \mathbb{R}$, on note $\Phi_x : \mathcal{C}_b(\mathbb{R}, \mathbb{R}) \rightarrow \mathbb{R}^{\mathbb{N}}$ l'application linéaire définie par

$$\Phi_x(f) = (f(x), f(x+1), \dots).$$

L'espace vectoriel \mathcal{M} des applications linéaires de $\mathbb{R}^{\mathbb{N}}$ dans lui-même s'identifie naturellement à l'espace vectoriel des matrices infinies indexées par $\mathbb{N} \times \mathbb{N}$. Pour tout $x \in \mathbb{R}$, et tout $f \in \mathcal{C}_b(\mathbb{R}, \mathbb{R})$, on a

$$\mathbf{L}(f)(x) = \mathbf{K}\Phi_x(f),$$

où \mathbf{K} est la matrice infinie définie pour tout i et j dans \mathbb{N} par

$$(\mathbf{K})_{i,j} := \lambda(\delta_{\{i+1=j\}} - \delta_{\{i=j\}}).$$

Cette formule n'est que l'écriture matricielle de la définition de \mathbf{L} . Soit \mathbf{I} et \mathbf{J} les matrices infinies définies pour tout i et j dans \mathbb{N} par $\mathbf{I}_{i,j} = \delta_{\{i=j\}}$ et $\mathbf{J}_{i,j} = -\delta_{\{i+1=j\}}$. Pour tout $t \geq 0$, on note $\mathbf{P}_t := \exp(t\mathbf{L})$ la matrice infinie obtenue par la série exponentielle classique. Montrer que pour $t > 0$

$$\mathbf{P}_t := e^{t\mathbf{L}} = e^{-\lambda t\mathbf{I}} e^{\lambda t\mathbf{J}} = e^{-\lambda t} \mathbf{Q}_{\lambda t},$$

où \mathbf{Q}_u est la matrice infinie donnée pour tout i et j dans \mathbb{N} par

$$(\mathbf{Q}_u)_{i,j} = \frac{u^{j-i}}{(j-i)!} \delta_{\{i \leq j\}}.$$

Soit $(N_t)_{t \geq 0}$ un processus de Poisson simple d'intensité λ . Pour tout $t \geq 0$, on désigne par $P_t : \mathcal{C}_b(\mathbb{R}, \mathbb{R}) \rightarrow \mathcal{C}_b(\mathbb{R}, \mathbb{R})$ l'application linéaire définie pour toute fonction $f \in \mathcal{C}_b(\mathbb{R}, \mathbb{R})$ et tout $x \in \mathbb{R}$ par

$$P_t(f)(x) := \mathbb{E}(f(x + N_t)).$$

Déduire de ce qui précède que pour tout $x \in \mathbb{R}$ et toute fonction $f \in \mathcal{C}_c(\mathbb{R}, \mathbb{R})$,

$$P_t \Phi_x(f) = P(f)(x).$$

Montrer plus généralement que cette formule reste valable pour $f \in \mathcal{C}_b(\mathbb{R}, \mathbb{R})$. Montrer que $P_0 = I$ et que pour tout $s \geq 0$ et $t \geq 0$, la propriété de semi-groupe suivante à lieu

$$P_{t+s} = P_t P_s = P_s P_t.$$

Réciproquement, soit $(N_t)_{t \geq 0}$ un processus de Poisson simple d'intensité λ . Montrer que pour toute fonction $f \in \mathcal{C}_c(\mathbb{R}, \mathbb{R})$ et tout $x \in \mathbb{R}$,

$$\frac{d}{dt} [P_t(f)(x)]_{t=0} = L(f)(x).$$

L'opérateur L est le *générateur infinitésimal de Markov* du processus de Poisson simple $(N_t)_{t \geq 0}$ d'intensité λ . La famille $(P_t)_{t \geq 0}$ est le *semi-groupe de Markov* associé au processus de Poisson simple $(N_t)_{t \geq 0}$ d'intensité λ .

Remarque 4.2 (Analogie pour les processus de Markov à temps discret & continu). Un processus de Markov est un processus qui satisfait à la propriété de Markov sur les lois conditionnelles « passé-présent-futur ». De manière générale, à un processus de Markov à temps continu $(X_t)_{t \geq 0}$ et à valeurs dans \mathbb{N} , on peut associer un semi-groupe de Markov $(P_t)_{t \geq 0}$ défini par $P_t(f)(x) = \mathbb{E}(f(X_t) | X_0 = x)$. Ce semi-groupe est noté pour abrégé $(\exp(tL))_{t \in \mathbb{N}}$, où L est le générateur infinitésimal de Markov défini par $L(f)(x) = \partial_{t=0} P_t(f)(x)$.

Si $(X_n)_{n \in \mathbb{N}}$ est une chaîne de Markov de matrice de transition T , alors $(T^n)_{n \in \mathbb{N}}$ est un semi-groupe de Markov à temps discret: $T^{n+m} = T^n T^m$ et $T^0 = I$. Ainsi, $(T^n)_{n \in \mathbb{N}}$ pour la chaîne de Markov et $(P_t)_{t \geq 0}$ pour un processus de Markov jouent le même rôle: ils décrivent la loi du processus le long du temps. La notion de matrice de transition en temps discret est remplacée en temps continu par la notion de générateur infinitésimal de Markov, mieux adaptée aux calculs. Le tableau analogique suivant résume ce parallèle.

Chaîne de Markov $(X_n)_{n \in \mathbb{N}}$ de matrice T	Processus de Markov $(X_t)_{t \geq 0}$ de générateur L
Temps discret $n \in \mathbb{N}$	Temps continu $t \in \mathbb{R}_+$
Matrice de transition T	$P_1 = \exp(L)$
$\mathcal{L}(X_n X_0) = T^n = \exp(n \log T)$	$\mathcal{L}(X_t X_0) = P_t = \exp(tL)$
$\log T$	Générateur infinitésimal de Markov L

Les chaînes de Markov sont les processus de Markov à *temps discret* les plus simples. Parmi les chaînes de Markov non triviales les plus simples, on distingue le processus de Bernoulli de paramètre $p \in]0, 1[$. Sa matrice infinie de transition T est donnée pour tout $(i, j) \in \mathbb{N} \times \mathbb{N}$ par $T_{i,j} := (1 - p)\delta_{\{i=j\}} + p\delta_{\{i+1=j\}} = (1 - p)I + pJ$. Le processus de Poisson simple est le processus de Markov à *temps continu* le plus simple. Les temps inter-saut d'un processus de Poisson simple sont i.i.d. et suivent des lois exponentielles, tandis que ceux d'un processus de Bernoulli sont i.i.d. et suivent des lois géométriques⁵. Ces deux processus correspondent au comptage de lois exponentielles et géométriques respectivement.

⁵A contrario, les temps inter-sauts d'une chaîne de Markov homogène de matrice de transition T ne sont pas indépendants ni équidistribués. Ils suivent cependant des lois géométriques puisque $\mathbb{P}(X_{n+1} \neq X_n | X_n = x) = \mathcal{G}(1 - T_{x,x})$. Il est alors clair que ces temps inter-sauts sont équidistribués si et seulement si les coefficients diagonaux de T sont tous identiques. De même, ils sont indépendants si et seulement si toutes les lignes de T sont identiques, et l'on dit alors que la chaîne est homogène en espace.

Processus de Bernoulli de paramètre $0 < p < 1$		Processus de Poisson simple d'intensité $\lambda > 0$
Discret $n \in \mathbb{N}$	Temps	Continu $t \in \mathbb{R}_+$
\mathbb{N}	Valeurs	\mathbb{N}
$-\log(1-p)$	Intensité	λ
Géométrique $\mathcal{G}(p)$	Inter-sauts	Exponentielle $\mathcal{E}(\lambda)$
Négative-binomiale (Pascal) $\mathcal{G}(p)^{*n}$	n -ième saut	Loi Gamma $\mathcal{E}(\lambda)^{*n}$
Binomiale $\mathcal{B}(n, p)$	Loi à temps fixe	Poisson $\mathcal{P}(t\lambda)$

Ici, $\mathcal{G}(p) := p \sum_{n=0}^{+\infty} (1-p)^n \delta_n$. Comme un processus de Bernoulli est indexé par un temps discret, il peut sauter consécutivement un nombre arbitraire de fois, et $\mathbb{P}(B_{n+k} = k + B_n) = p^k$. Ce phénomène n'est pas possible pour le processus de Poisson simple puisque les temps de sauts suivent une loi à densité (exponentielle) par rapport à la mesure de Lebesgue, qui ne charge donc pas le point $\{0\}$.

L'approche « générateur infinitésimal de Markov » pour le processus de Poisson simple s'avère très utile pour l'étude des files d'attente. Ainsi, la longueur d'une file d'attente M/M/1 avec intensité d'arrivée des clients $\lambda > 0$ et intensité de service $\mu > 0$ est un processus de Markov $(F_t)_{t \geq 0}$ à temps continu et à valeurs dans \mathbb{N} , tout comme le processus de Poisson simple. Son générateur infinitésimal $\mathbf{L}_{\lambda, \mu}$ est donné par

$$\begin{aligned} \mathbf{L}_{\lambda, \mu}(f)(x) &:= \lambda(f(x+1) - f(x)) + \mu(f(x-1) - f(x)) \\ &= \lambda f(x+1) + (\lambda + \mu) f(x) + \mu f(x-1). \end{aligned}$$

Les temps inter-saut de la file d'attente sont i.i.d. de loi géométrique $\mathcal{G}(\lambda + \mu)$. Lorsque $\lambda < \mu$, la file d'attente atteint un état d'équilibre lorsque le temps tend vers l'infini (ergodicité), quel que soit son point de départ. Cet équilibre est caractérisé par une loi de probabilité invariante unique. Une mesure η sur \mathbb{N} est invariante si et seulement si elle vérifie $\mathbb{E}_\eta(\mathbf{L}_{\lambda, \mu} f) = 0$ pour toute fonction f . En d'autres termes, on cherche une suite $(\eta(n))_{n \in \mathbb{N}}$ telle que pour toute suite $(f(n))_{n \in \mathbb{N}}$,

$$\sum_{n=0}^{+\infty} \eta(n) [\lambda f(n+1) + (\lambda + \mu) f(n) + \mu f(n-1)] = 0,$$

avec la convention $f(-1) = 0$. On montre alors, en considérant des suites $(f(n))_{n \in \mathbb{N}}$ bien choisies, que la loi géométrique $\mathcal{G}(\lambda/\mu)$ est invariante lorsque $\lambda < \mu$. On notera que la mesure invariante du processus de Poisson simple ($\mu = 0$) n'est pas unique, et tout multiple de la mesure uniforme (i.e. de comptage) sur \mathbb{N} convient, et elles ne se normalisent pas en une loi de probabilité. La théorie actuelle des processus de Markov dépasse le cadre restreint de cette remarque, mais repose essentiellement sur les intuitions présentées ici.

Exercice 4.3 (Concentration poissonnienne et inégalité de Sobolev logarithmique). On note π_λ la loi de Poisson $\mathcal{P}(\lambda)$. Soit $(N_t)_{t \geq 0}$ un processus de Poisson simple d'intensité $\lambda > 0$. Pour tout t , on note P_t l'opérateur linéaire agissant sur les éléments de $\mathcal{C}_b(\mathbb{R}, \mathbb{R})$ par $P_t^x(f) := P_t(f)(x) := \mathbb{E}_{P_t^x}(f) = \mathbb{E}(f(x + N_t))$. On peut également voir P_t^x comme la loi de probabilité $\mathcal{L}(x + N_t) = \delta_x * \pi_\lambda$. Montrer que la famille d'opérateurs $(P_t)_{t \geq 0}$ forme un semi-groupe de contractions pour la composition des opérateurs, d'élément neutre P_0 , et vérifiant $P_t(1) = 1$. Un tel semi-groupe est qualifié de « markovien ». Montrer que pour tout $f \in \mathcal{C}_b(\mathbb{R}, \mathbb{R})$, tout $t > 0$

$$\mathbf{L}f := \partial_t P_t f = \lambda Df,$$

où $(Df)(x) := f(x+1) - f(x)$. En utilisant l'indépendance des accroissements de $(N_t)_{t \geq 0}$, montrer que les opérateurs \mathbf{L} , P_t et D commutent pour tout $t \geq 0$. Soit Φ la fonction définie sur \mathbb{R}_+ par $\Phi(u) := u \log u$ et soit Ψ la fonction définie pour tout couple $(u, v) \in \mathbb{R}^2$ avec $u + v \geq 0$ et $u \geq 0$ par

$$\Psi(u, v) := \Phi(u + v) - \Phi(u) - \Phi'(u)v.$$

Montrer que Φ et Ψ sont strictement convexes sur leur domaine de définition. Montrer que pour toute $g \in \mathcal{C}_b(\mathbb{R}, \mathbb{R}_+)$ et tout $s > 0$

$$\partial_s P_t(\Phi(g)) = \lambda P_s(\Psi(g, Dg)).$$

En utilisant la propriété de semi-groupe $P_t = P_{t-s} \circ (P_s)$ pour $0 \leq s \leq t$, montrer que pour toute $f \in \mathcal{C}_b(\mathbb{R}, \mathbb{R})$ et tout $t \geq 0$

$$P_t(\Phi(f)) - \Phi(P_t(f)) = \int_0^t P_s(\Psi(g, Dg)) ds,$$

où $g := P_{t-s}(f)$. En utilisant l'inégalité de Jensen pour la fonction convexe Ψ et les propriétés de commutations établies plus haut, montrer que pour toute $f \in \mathcal{C}_b(\mathbb{R}, \mathbb{R}_+)$ et tout $t \geq 0$

$$P_t(\Phi(f)) - \Phi(P_t(f)) \leq \lambda P_t(\Psi(f, Df)).$$

Une telle inégalité est dite « de Sobolev logarithmique » ou encore « de Gross ». Montrer que pour tout (u, v) avec $u \geq 0$ et $u + v \geq 0$, $\Psi(u, v) \leq v^2/u$. On a coutume d'appeler « entropie » de f pour la loi π_λ la quantité

$$\mathbf{ent}_{\pi_\lambda}(f) := \mathbb{E}_{\pi_\lambda}(\Phi(f)) - \Phi(\mathbb{E}_{\pi_\lambda}(f)).$$

En remarquant que $P_1 = \pi_\lambda$, en déduire que pour toute $f \in \mathcal{C}_b(\mathbb{R}, \mathbb{R}_+^*)$

$$\mathbf{ent}_{\pi_\lambda}(f) \leq \lambda \mathbb{E}_{\pi_\lambda} \left(\frac{|Df|^2}{f} \right).$$

Montrer que la constante λ devant le membre de droite est optimale en considérant une fonction f de la forme $f(x) = \exp(-\alpha x)$ avec α bien choisi. En appliquant l'inégalité précédente à la fonction $f(x) = \exp(\mu F)$ où F est une fonction bornée continue qui vérifie $|DF| \leq 1$, et en déduire que

$$\forall \mu \geq 0, K'(\mu) \leq \lambda \exp(2\mu),$$

où $K(\mu) := \mu^{-1} \log \mathbb{E}(\exp(\mu F))$ avec $K'(0) = \mathbb{E}_{\pi_\lambda}(F)$. En déduire que pour tout $\mu \geq 0$,

$$\log \mathbb{E}_{\pi_\lambda}(\exp(\mu F)) \leq \mu \mathbb{E}_{\pi_\lambda}(F) + \mu \lambda (\exp(2\mu) - 1).$$

Étendre cette inégalité au cas où F n'est pas bornée en approximant F par la suite de fonctions bornées $F_n := \max(-n, \min(n, F))$ qui vérifie $|DF_n| \leq |DF|$. En déduire par l'inégalité de Tchebychev que pour tout $r > 0$

$$\pi_\lambda(x \in \mathbb{R}, |F(x) - \mathbb{E}_{\pi_\lambda}(F)| \geq r) \leq 2 \exp \left(-\frac{r}{4} \log \left(1 + \frac{r}{2\lambda} \right) \right).$$

Ce type d'inégalité exprime que la loi de F sous π_λ se « concentre » exponentiellement autour de sa moyenne. Elle quantifie la propriété de tension pour les lois de probabilité. Soit X_1, \dots, X_n des v.a. i.i.d. de loi de Poisson π_λ . En utilisant le fait que $\mathcal{L}(X_1 + \dots + X_n) = (\pi_\lambda)^{*n} = \pi_{n\lambda}$, déduire de ce qui précède que pour tout $r > 0$

$$\mathbb{P} \left(\left| \frac{X_1 + \dots + X_n}{n} - \lambda \right| \geq r \right) \leq 2 \exp \left(-\frac{rn}{4} \log \left(1 + \frac{r}{2\lambda} \right) \right).$$

L'inégalité obtenue a le grand mérite de ne pas être asymptotique. Comparer ce résultat à ce que donne le théorème de Cramér (Principe de Grandes Déviations) et le théorème central limite (approximation gaussienne).