Projet GM3

Processus de Poisson et temps d'attente

Elliot Sisteron
Antoni Markovski

À l'attention de
M. Lescot

Introduction

Dans notre société actuelle, où tout se réalise de plus en plus vite, la frustration de l'attente est une sensation des plus déplaisante. Une manière de moins redouter cette attente est d'avoir un certain contrôle sur celle-ci. En effet, de nombreuses études ont démontré que le sentiment d'impatience était très fortement lié à notre capacité à quantifier l'attente dans une situation donnée.

Dans ce projet, on se propose de résoudre un problème de ce type : nous allons essayer de déterminer le temps moyen d'attente nécessaire à un passager avant l'arrivée de son bus. Pour cela, on sait que la durée moyenne écoulée entre le passage de deux bus est de 15 minutes. On notera par la suite :

— $(T_n)_{n\geq 0}$ la suite de variables aléatoires réelles représentant l'horaire (exprimée en minutes) de passage de l'autobus numéro n. On pose $T_0 = 0$.

— $(X_n)_{n\geq 0} = (T_{n+1} - T_n)_{n\geq 0}$ la suite de variables aléatoires qui exprime l'écart temporel entre deux passages de bus adjacents. On supposera que cette suite est constituée de variables aléatoires *i.i.d* (indépendantes et identiquement distribuées) de loi exponentielle de paramètre λ à déterminer.

— A_t la variable aléatoire représentant l'attente nécessaire au passager pour voir un autobus passer, sachant que cet usager de la ligne est arrivé à l'instant $t > 0$ fixé.

Le but sera donc ici de déterminer $\mathbb{E}[A_t]$.

Nous commencerons par présenter le processus de Poisson pour pouvoir mieux aborder le problème. Dès lors, cela nous aidera à déterminer la constante λ recherchée. Puis, nous nous intéresserons à la simulation de ce problème pour conjecturer le résultat attendu. Enfin, nous calculerons rigoureusement le temps d'attente moyen pour le comparer avec le résultat de la simulation précédente.

Table des matières

Introduction **2**

1 Présentation du processus de Poisson **4**
- 1.1 Modélisation de l'expérience aléatoire . . . 4
- 1.2 Description d'un processus de Poisson . . . 5
  - 1.2.1 Qui est Poisson ? . . . 5
  - 1.2.2 Caractérisation du processus de Poisson par un processus de renouvellement 5
  - 1.2.3 Processus de comptage . . . 6
- 1.3 Détermination de λ . . . 10
  - 1.3.1 Calcul de l'espérance d'une variable aléatoire suivant une loi exponentielle 10
  - 1.3.2 Calcul de λ . . . 10

2 Simulation et estimation de $\mathbb{E}[A_t]$ **11**
- 2.1 Comment simuler une variable aléatoire réelle ? . . . 11
  - 2.1.1 Simulation d'une variable aléatoire exponentielle . . . 12
- 2.2 Conjecture préliminaire . . . 12
- 2.3 Simulation de A_t . . . 13
  - 2.3.1 Algorithme de simulation . . . 13
  - 2.3.2 Implémentation de l'algorithme en langage C . . . 14
- 2.4 Approximation et conjecture de $\mathbb{E}[A_t]$ . . . 16

3 Calcul de l'espérance de A_t **18**
- 3.1 Détermination de la loi de A_t . . . 18
- 3.2 Calcul de $\mathbb{E}[A_t]$ . . . 19
- 3.3 Comparaison avec les résultats précédents . . . 20

Conclusion **23**

A La variable aléatoire de comptage dans un intervalle $N_{]a;b]}$ **24**

B Une autre manière de calculer I_n **25**

C Une autre façon de déterminer la loi de A_t et de B_t **27**

D Programme de simulation **29**

Bibliographie **33**

Présentation du processus de Poisson

Nous allons ici commencer par modéliser graphiquement les *v.a* de l'énoncé pour bien comprendre le problème. Nous présenterons ensuite rapidement la vie de Siméon Denis Poisson. Enfin, nous utiliserons une approche spécifique pour définir rigoureusement le processus de Poisson avant de déterminer λ, la constante recherchée.

1.1 Modélisation de l'expérience aléatoire

On est sur un espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$. On s'intéresse ici aux variables aléatoires de temps (donc réelles) $(X_n)_{n\geq 0}$ et $(T_n)_{n\geq 0}$. Toutes les durées sont exprimées en minutes.
$\forall n \in \mathbb{N}$, on a :

$$\begin{array}{rcccc} T_n & : & \Omega & \rightarrow & \mathbb{R}_+ \\ & & \omega & \mapsto & T_n(\omega) \end{array}$$

et,

$$\begin{array}{rcccc} X_n & : & \Omega & \rightarrow & \mathbb{R}_+ \\ & & \omega & \mapsto & T_{n+1}(\omega) - T_n(\omega) \end{array}$$

On réalise un petit schéma pour mieux visualiser les variables aléatoires mises en jeu, les X_n représentent le temps séparant deux autobus et les T_n évaluent l'heure de passage de chaque autobus à l'arrêt :

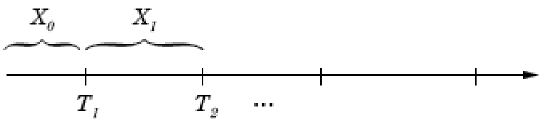

Figure 1.1 – Un schéma pour mieux visualiser nos v.a

On fixe $t > 0$. On s'intéresse à la variable aléatoire A_t du temps d'attente d'un usager à l'arrêt.

$$\begin{array}{rcccc} A_t & : & \Omega & \rightarrow & \mathbb{R}_+ \\ & & \omega & \mapsto & A_t(\omega) \end{array}$$

Le but sera ici de déterminer la loi de A_t pour en déduire son espérance.

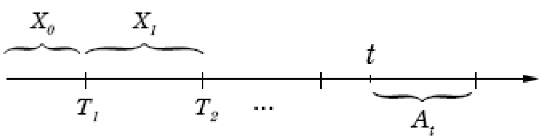

Figure 1.2 – Le même schéma avec A_t

1.2 Description d'un processus de Poisson

1.2.1 Qui est Poisson ?

Siméon Denis Poisson est un mathématicien Français du XVIII^{e} – XIX^{e} siècle. Sa famille voulait qu'il devint médecin, et le poussa à faire des études dans ce domaine. Il abandonna cette idée en 1798, pour aller étudier les mathématiques à l'École polytechnique où il fit ses études auprès de grands noms des mathématiques, notamment P. Laplace et J. Lagrange. Il devint ensuite enseignant à l'Ecole polytechnique en 1802. La phrase culte de ce mathématicien de renom est : « La vie n'est bonne qu'à deux choses : découvrir des mathématiques et enseigner les mathématiques ! ».

Poisson est connu pour ses nombreuses publications en mathématiques et en sciences physiques. Parmi ses publications en mathématiques pures figurent notamment une série d'articles sur les intégrales définies, et ses recherches sur les séries de Fourier ont annoncé celles de Dirichlet et de Riemann sur ce sujet. Il s'intéressait beaucoup aux problèmes aléatoires. Le processus de Poisson est l'un des processus aléatoires les plus importants dans la théorie des probabilités. Il est largement utilisé pour modéliser des points aléatoires dans le temps et l'espace, tels que les temps d'émissions radioactives, les temps d'arrivée des clients dans un centre de service, et la recherche de défauts dans des matériaux.

Nous allons voir que plusieurs distributions de probabilités importantes apparaissent naturellement dans le processus de Poisson - la distribution de Poisson (d'où le nom donné au processus), la loi exponentielle, et la distribution gamma. Le processus a une belle structure mathématique, et est utilisé comme une base pour la construction d'un certain nombre d'autres processus aléatoires, plus complexes. Dans notre cas, on va l'utiliser pour représenter des passages de bus à un arrêt.

1.2.2 Caractérisation du processus de Poisson par un processus de renouvellement

On dira qu'une suite de variables aléatoires $(T_n)_{n\geq 0}$ est un processus de renouvellement lorsque la suite définie par $(X_n)_{n\geq 0} = (T_{n+1} - T_n)_{n\geq 0}$ est constituée de variable aléatoires *i.i.d.* Si, de plus, les X_n suivent une loi exponentielle de paramètre λ, on dira que les T_n définissent un processus de Poisson d'« intensité » λ. Dans notre cas, on est bien face à un processus de Poisson.

Dans un processus de Poisson, on a donc que : $\forall n \geq 0,\ X_n \hookrightarrow \xi(\lambda)$. La densité de probabilité de X_n est ainsi définie :

$$\forall x \in \mathbb{R},\ f_{X_n}(x) = \lambda e^{-\lambda x} \mathbb{1}_{\mathbb{R}_+}(x)$$

On peut donc en déduire la fonction de répartition F_{X_n} de chacune de ces variables aléatoires

$$\begin{aligned}\forall x \in \mathbb{R},\ F_{X_n}(x) &= \mathbb{P}(X_n \leq x)\\ &= \int_{-\infty}^{x} f_{X_n}(t)dt\\ &= \int_{-\infty}^{x} \lambda e^{-\lambda t}\mathbb{1}_{\mathbb{R}_+}(t)dt\\ &= \int_{0}^{x} \lambda e^{-\lambda t}dt\\ &= \lambda\left[-\frac{1}{\lambda}e^{-\lambda t}\right]_0^x\\ &= \left(1-e^{-\lambda x}\right)\mathbb{1}_{\mathbb{R}_+}(x)\end{aligned}$$

1.2.3 Processus de comptage

Une autre manière de définir le processus de Poisson serait à partir de N_t (pour $t > 0$ fixé), la variable aléatoire qui « compte » le nombre de T_n déjà produits à l'instant t. Autrement dit,

$$N_t = \sum_{n\geq 0} \mathbb{1}_{\{T_n \leq t\}}$$

$$\begin{array}{cccc} N_t & : & \Omega & \to & \mathbb{N} \\ & & \omega & \mapsto & \sum_{n\geq 0}\mathbb{1}_{\{T_n\leq t\}}(\omega)\end{array}$$

On a donc la relation suivante :

$$\begin{aligned}\forall \omega \in \Omega,\ \forall n \in \mathbb{N},\ N_t(\omega) = n &\Leftrightarrow \forall k \in [\![1;n]\!],\ \mathbb{1}_{\{T_k\leq t\}}(\omega) = 1\ et\ \forall k > n,\ \mathbb{1}_{\{T_k\leq t\}}(\omega) = 0\\ &\Leftrightarrow \forall k \in [\![1;n]\!],\ T_k(\omega) \leq t\ et\ \forall k > n,\ T_k(\omega) > t\\ &\Leftrightarrow T_n(\omega) \leq t < T_{n+1}(\omega)\end{aligned}$$

Ici, la variable aléatoire N_t compte le nombre de bus passés dans l'intervalle de temps $[0;t]$. On aurait aussi pu s'intéresser à la variable aléatoire de comptage dans un intervalle quelconque $N_{]a;b]}$ [1].

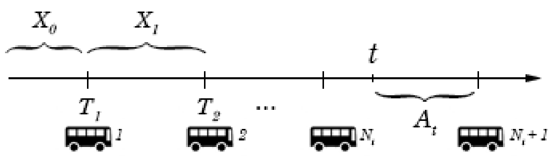

FIGURE 1.3 – Un schéma pour visualiser N_t

On ne va pas caractériser le processus de Poisson avec N_t, mais nous allons tout de même étudier sa loi à partir de la loi des T_n. En effet :

$$\begin{aligned}\omega \in \{N_t = n\} &\Leftrightarrow \omega \in \{T_n \leq t < T_{n+1}\}\\ &\Leftrightarrow \omega \in \{T_n \leq t\} \cap \{t < T_{n+1}\}\end{aligned}$$

1. Étudiée dans l'annexe A

Donc,

$$\begin{aligned}
\mathbb{P}(N_t = n) &= \mathbb{P}(T_n \leq t \cap t < T_{n+1}) \\
&= \mathbb{P}(T_n \leq t) + \mathbb{P}(T_{n+1} > t) - \mathbb{P}(T_n \leq t \cup t < T_{n+1}) \\
&= F_{T_n}(t) + 1 - F_{T_{n+1}}(t) - 1
\end{aligned}$$

Car

$$\begin{aligned}
\omega \in \{T_n \leq t \cup t < T_{n+1}\} &\Leftrightarrow \omega \in \{T_n \leq t\} \; ou \; \omega \in \{t < T_{n+1}\} \\
&\Leftrightarrow \omega \in \Omega
\end{aligned}$$

D'où :

$$\boxed{\forall n \in \mathbb{N}, \; \mathbb{P}(N_t = n) = F_{T_n}(t) - F_{T_{n+1}}(t)}$$

Il nous reste à trouver la loi de des T_n. On remarque pour cela que, par somme télescopique :

$$\forall n \in \mathbb{N}^*, \; T_n = \sum_{k=0}^{n-1} X_k$$

Or, on rappelle que pour deux variables aléatoires réelles X et Y indépendantes admettant des densités de probabilités f_X et f_Y, la densité de leur somme f_{X+Y} est donnée par le produit de convolution :

$$\forall x \in \mathbb{R}, \; f_{X+Y}(x) = (f_X * f_Y)(x) = \int_{-\infty}^{\infty} f_X(t) f_Y(x-t) dt$$

On peut donc retrouver la densité de probabilité de T_n par récurrence. $T_1 = X_0 \hookrightarrow \xi(\lambda)$, donc :

$$f_{T_1}(x) = f_{X_0}(x) = \lambda e^{-\lambda x} \mathbb{1}_{\mathbb{R}_+}(x)$$

$T_2 = X_0 + X_1$, ainsi :

$$\begin{aligned}
f_{T_2}(x) = f_{X_0+X_1}(x) &= \int_{-\infty}^{\infty} f_{X_0}(t) f_{X_1}(x-t) dt \\
&= \int_{-\infty}^{\infty} \lambda e^{-\lambda t} \mathbb{1}_{\mathbb{R}_+}(t) \lambda e^{-\lambda(x-t)} \mathbb{1}_{\mathbb{R}_+}(x-t) dt \\
&= \lambda^2 e^{-\lambda x} \int_0^x dt \\
&= \lambda^2 x e^{-\lambda x} \mathbb{1}_{\mathbb{R}_+}(x)
\end{aligned}$$

$T_3 = T_2 + X_2$, d'où :

$$\begin{aligned}
f_{T_3}(x) = f_{T_2+X_2}(x) &= \int_{-\infty}^{\infty} f_{T_2}(t) f_{X_2}(x-t) dt \\
&= \int_{-\infty}^{\infty} \lambda^2 t e^{-\lambda t} \mathbb{1}_{\mathbb{R}_+}(t) \lambda e^{-\lambda(x-t)} \mathbb{1}_{\mathbb{R}_+}(x-t) dt \\
&= \lambda^3 e^{-\lambda x} \int_0^x t dt \\
&= \lambda^3 \frac{x^2}{2} e^{-\lambda x} \mathbb{1}_{\mathbb{R}_+}(x)
\end{aligned}$$

On peut ainsi conjecturer que :

$$\boxed{\forall n \in \mathbb{N}^*, \; \forall x \in \mathbb{R}, \; f_{T_n}(x) = \lambda^n \frac{x^{n-1}}{(n-1)!} e^{-\lambda x} \mathbb{1}_{\mathbb{R}_+}(x)}$$

L'initialisation est déjà faite, il suffit de montrer l'hérédité. Supposons la propriété correcte au rang $n \in \mathbb{N}^*$ fixé.
$T_{n+1} = T_n + X_n$, donc :

$$\begin{aligned} f_{T_{n+1}}(x) = f_{T_n+X_n}(x) &= \int_{-\infty}^{\infty} f_{T_n}(t) f_{X_n}(x-t)dt \\ &= \int_{-\infty}^{\infty} \lambda^n \frac{t^{n-1}}{(n-1)!} e^{-\lambda t} \mathbb{1}_{\mathbb{R}_+}(t) \lambda e^{-\lambda(x-t)} \mathbb{1}_{\mathbb{R}_+}(x-t)dt \\ &= \frac{\lambda^{n+1}}{(n-1)!} e^{-\lambda x} \int_0^x t^{n-1} dt \\ &= \lambda^{n+1} \frac{x^n}{n!} e^{-\lambda x} \mathbb{1}_{\mathbb{R}_+}(x) \end{aligned}$$

Ainsi, on peut en déduire que la conjecture précédente est vraie. On remarque au passage que $T_n \hookrightarrow \Gamma(n, \lambda)$. On aurait pu s'attendre à ce résultat, vu que la somme de n variables exponentielles $i.i.d$ de paramètre λ suit une loi gamma de paramètre n et λ.

Pour en déduire la loi de T_n, il nous reste maintenant à étudier la fonction de répartition de T_n :

$$\begin{aligned} \forall n \in \mathbb{N}^*,\ \forall x \in \mathbb{R},\ F_{T_n}(x) &= \mathbb{P}(T_n \leq x) \\ &= \int_{-\infty}^{x} \lambda^n \frac{t^{n-1}}{(n-1)!} e^{-\lambda t} \mathbb{1}_{\mathbb{R}_+}(t)dt \\ &= \frac{\lambda^n}{(n-1)!} \int_0^x t^{n-1} e^{-\lambda t} dt \end{aligned}$$

On pose :

$$I_n(x) = \int_0^x t^{n-1} e^{-\lambda t} dt$$

On va encore utiliser une récurrence[2] sur n pour calculer I_n. Commençons par conjecturer le résultat en utilisant une intégration par partie, on fixe $x \in \mathbb{R}_+$ et $n \in \mathbb{N}^*$:

$$\begin{aligned} I_n(x) &= \int_0^x t^{n-1} e^{-\lambda t} dt \\ &= \left[-\frac{t^{n-1}}{\lambda} e^{-\lambda t} \right]_0^x - \int_0^x -(n-1) \frac{t^{n-2}}{\lambda} e^{-\lambda t} dt \\ &= -\frac{x^{n-1}}{\lambda} e^{-\lambda x} - (n-1) \frac{x^{n-2}}{\lambda^2} e^{-\lambda x} - ... - \frac{(n-1)!}{1!} \frac{x}{\lambda^{n-1}} e^{-\lambda x} - \frac{(n-1)!}{\lambda^n} e^{-\lambda x} + \frac{(n-1)!}{\lambda^n} \end{aligned}$$

On se doute donc que I_n sera de la forme :

$$\begin{aligned} I_n(x) &= \frac{(n-1)!}{\lambda^n} - e^{-\lambda x} \sum_{k=1}^{n} \frac{(n-1)!}{(n-k)!} \frac{x^{n-k}}{\lambda^k} \\ &= \frac{(n-1)!}{\lambda^n} - e^{-\lambda x} \sum_{k=0}^{n-1} \frac{(n-1)!}{k!} \frac{(\lambda x)^k}{\lambda^n} \end{aligned}$$

2. Voir l'annexe B pour une autre démonstration plus élégante

En effet, pour ce qui est de l'initialisation, à $x \in \mathbb{R}_+$ fixé :

$$\begin{aligned}
I_1(x) &= \int_0^x e^{-\lambda t} dt \\
&= \frac{1}{\lambda}(1 - e^{-\lambda x}) \\
&= \frac{(1-1)!}{\lambda^1} - e^{-\lambda x} \frac{(1-1)!}{0!} \frac{(\lambda x)^0}{\lambda^1}
\end{aligned}$$

Ce qui vérifie la propriété au rang 1. On s'occupe de l'hérédité, supposons la propriété au rang n, calculons I_{n+1} (toujours à $x \in \mathbb{R}_+$ fixé) :

$$\begin{aligned}
I_{n+1}(x) &= \int_0^x t^n e^{-\lambda t} dt \\
&= -\frac{1}{\lambda}\left[t^n e^{-\lambda t}\right]_0^x + \frac{n}{\lambda} I_n \\
&= -\frac{x^n}{\lambda} e^{-\lambda x} + \frac{n}{\lambda} \frac{(n-1)!}{\lambda^n} - \frac{n}{\lambda} e^{-\lambda x} \sum_{k=0}^{n-1} \frac{(n-1)!}{k!} \frac{(\lambda x)^k}{\lambda^n} \\
&= \frac{n!}{\lambda^{n+1}} - e^{-\lambda x} \sum_{k=0}^{n} \frac{n!}{k!} \frac{(\lambda x)^k}{\lambda^{n+1}}
\end{aligned}$$

Finalement, on a bien :

$$\boxed{\forall n \in \mathbb{N}^*,\ \forall x \in \mathbb{R},\ I_n(x) = \left(\frac{(n-1)!}{\lambda^n} - e^{-\lambda x} \sum_{k=0}^{n-1} \frac{(n-1)!}{k!} \frac{(\lambda x)^k}{\lambda^n} \right) \mathbb{1}_{\mathbb{R}_+}(x)}$$

D'où :

$$\begin{aligned}
\forall n \in \mathbb{N}^*, F_{T_n}(x) &= \frac{\lambda^n}{(n-1)!} I_n(x) \\
&= \left(1 - e^{-\lambda x} \sum_{k=0}^{n-1} \frac{1}{k!} \frac{(\lambda x)^k}{1} \right) \mathbb{1}_{\mathbb{R}_+}(x)
\end{aligned}$$

N'oublions pas le cas où $n = 0$,

$$F_{T_0}(x) = \mathbb{P}(T_0 \leq x) = \mathbb{P}(0 \leq x) = \mathbb{1}_{\mathbb{R}_+}(x)$$

Donc la loi de T_n est bien définie par :

$$\boxed{\forall n \in \mathbb{N},\ \forall x \in \mathbb{R},\ F_{T_n}(x) = \left(1 - e^{-\lambda x} \sum_{k=0}^{n-1} \frac{(\lambda x)^k}{k!} \right) \mathbb{1}_{\mathbb{R}_+}(x)}$$

Et finalement, il en découle la loi de N_t :

$$\begin{aligned}
\mathbb{P}(N_t = n) &= F_{T_n}(t) - F_{T_{n+1}}(t) \\
&= 1 - e^{-\lambda t} \sum_{k=0}^{n-1} \frac{(\lambda t)^k}{k!} - 1 + e^{-\lambda t} \sum_{k=0}^{n} \frac{(\lambda t)^k}{k!} \\
&= e^{-\lambda t} \frac{(\lambda t)^n}{n!}
\end{aligned}$$

Ainsi,

$$\boxed{\forall n \in \mathbb{N},\ \mathbb{P}(N_t = n) = e^{-\lambda t}\frac{(\lambda t)^n}{n!}}$$

On remarque que N_t suit une loi de Poisson de paramètre λt, on comprend alors mieux l'appellation « processus de Poisson » : $N_t \hookrightarrow \mathcal{P}(\lambda t)$.

1.3 Détermination de λ

On sait que la durée moyenne entre deux passages de bus est de 15 minutes. Ainsi, on peut en déduire que :

$$\forall n \geq 0,\ \mathbb{E}[X_n] = 15$$

Or, on sait que chaque $X_n \hookrightarrow \xi(\lambda)$. Nous avons donc besoin de calculer l'espérance d'une variable aléatoire de loi exponentielle.

1.3.1 Calcul de l'espérance d'une variable aléatoire suivant une loi exponentielle

Soit X une variable aléatoire réelle sur un espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$ telle que $X \hookrightarrow \xi(\lambda)$. Soit f_X sa densité de probabilité.

Par le théorème de transfert, on a que :
$$\begin{aligned}\mathbb{E}[X] &= \int_{-\infty}^{\infty} t f_X(t)dt \\ &= \int_{-\infty}^{\infty} t\lambda e^{-\lambda t}\mathbb{1}_{\mathbb{R}_+}(t)dt \\ &= \lambda\int_{0}^{\infty} te^{-\lambda t}dt\end{aligned}$$

Par une intégration par partie, il en résulte que :
$$\begin{aligned}\mathbb{E}[X] &= \lambda\left(\left[-\frac{t}{\lambda}e^{-\lambda t}\right]_0^{\infty} - \int_0^{\infty} -\frac{1}{\lambda}e^{-\lambda t}dt\right) \\ &= \lambda\left[-\frac{1}{\lambda^2}e^{-\lambda t}\right]_0^{\infty} \\ &= \frac{1}{\lambda}\end{aligned}$$

D'où :

$$\boxed{\mathbb{E}[X] = \frac{1}{\lambda}}$$

1.3.2 Calcul de λ

On peut ainsi en déduire λ :

$$\begin{aligned}\forall n \geq 0,\ \mathbb{E}[X_n] = 15 &\Leftrightarrow \frac{1}{\lambda} = 15 \\ &\Leftrightarrow \lambda = \frac{1}{15}\end{aligned}$$

D'où :

$$\boxed{\lambda = \frac{1}{15}}$$

Simulation et estimation de $\mathbb{E}[A_t]$

Ici, on s'intéressera d'abord à la simulation de variables aléatoires réelles, puis à la simulation de $v.a$ exponentielles pour pouvoir enfin simuler A_t. On émettra d'abord une conjecture préliminaire concernant son espérance, puis nous la simulerons pour en déduire une conjecture plus affinée.

2.1 Comment simuler une variable aléatoire réelle ?

Soit X une variable aléatoire réelle sur un espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$. On connaît la loi de X, c'est-à-dire la fonction de répartition F_X telle que $\forall x \in \mathbb{R},\ F_X(x) = \mathbb{P}(X \leq x)$. On cherche ici à simuler la $v.a.r$ X sur $(\Omega, \mathcal{F}, \mathbb{P})$. Pour cela, on supposera que l'on est déjà capable de simuler la loi uniforme $\mathcal{U}([0;1])$ (par exemple, la fonction $rand()$ en C est capable de simuler une telle loi).

On pose G_X l'application telle que $\forall u \in [0;1[,\ G_X(u) = \inf\{t \in \mathbb{R}\ /\ F_X(t) > u\}$ aussi appelée « inverse généralisée de F_X ». En effet, pour F_X continue et strictement croissante, on a $G_X = F_X^{-1}$. On remarque que cette fonction est croissante et *càdlàg* (c'est-à-dire continue à droite et limitée à gauche) du fait que F_X est croissante et *càdlàg*.

Théorème. *Si $U \hookrightarrow \mathcal{U}([0;1])$, $G_X(U)$ a la même loi que X.*

Preuve. *Calculons la f.d.r $F_{G_X(U)}$ de $G_X(U)$. Soit $x \in \mathbb{R}$:*

$$F_{G_X(U)}(x) = \mathbb{P}(G_X(U) \leq x)$$

Or, $\forall u \in [0;1[$

$$\begin{aligned} x < G_X(u) &\Rightarrow x < \inf\{t \in \mathbb{R}\ /\ F_X(t) > u\} \\ &\Rightarrow F_X(x) \leq u \end{aligned}$$

Donc, par contraposée,

$$\boxed{F_X(x) > u \Rightarrow x \geq G_X(u)}$$

Et, du fait de la croissance de F_X,

$$x \geq G_X(u) \Rightarrow F_X(x) \geq F_X(G_X(u))$$

Or, comme F_X est càdlàg, il existe $(t_n)_n$ une suite de réels décroissante telle que $\forall n \in \mathbb{N},\ F_X(t_n) > u$ et $t_n \underset{n\to+\infty}{\longrightarrow} G_X(u)$. Ainsi, $F_X(G_X(u)) \geq u$. D'où :

$$\boxed{x \geq G_X(u) \Rightarrow F_X(x) \geq u}$$

En revenant au calcul de la fonction de répartition précédente, on peut ainsi en déduire :

$$\begin{aligned} F_{G_X(U)}(x) &= \mathbb{P}(G_X(U) \leq x) \\ &= \mathbb{P}(U < F_X(x)) \end{aligned}$$

Or, la f.d.r de U est continue, donc $\mathbb{P}(U = F_X(x)) = 0$, donc $\mathbb{P}(U < F_X(x)) = \mathbb{P}(U \leq F_X(x))$. D'où :

$$\begin{aligned} F_{G_X(U)}(x) &= \mathbb{P}(U \leq F_X(x)) \\ &= F_U(F_X(x)) \\ &= F_X(x) \end{aligned}$$

car $F_X(x)$ est à valeurs dans $[0;1]$.

Ainsi par le théorème de caractérisation d'une loi par sa f.d.r, on obtient que la loi de X est la même que la loi de $G_X(U)$.

D'après ce théorème, on peut donc simuler n'importe quelle variable aléatoire réelle en trouvant son inverse généralisée.

2.1.1 Simulation d'une variable aléatoire exponentielle

On a vu que pour simuler une *v.a* exponentielle il fallait trouver son inverse généralisée. Or, ici, on remarque que la *f.d.r* de notre loi exponentielle est bijective donc $G_X = F_X^{-1}$. Trouvons cette fonction réciproque :

$$\begin{aligned} \forall x \in \mathbb{R}_+,\ F_X(x) = u &\Leftrightarrow 1 - e^{-\lambda x} = u \\ &\Leftrightarrow e^{-\lambda x} = 1 - u \\ &\Leftrightarrow x = -\frac{ln(1-u)}{\lambda} \end{aligned}$$

Donc,

$$F_X^{-1}(u) = -\frac{ln(1-u)}{\lambda}$$

Finalement, pour simuler la loi d'un *v.a* exponentielle, il suffit d'appliquer $F_X^{-1}(u)$ sur une distribution de $u \in [0;1[$ uniforme. On peut même prendre :

$$\boxed{\forall u \in]0;1],\ F_X^{-1}(u) = -\frac{ln(u)}{\lambda}}$$

ce qui revient exactement au même et facilite les calculs.

2.2 Conjecture préliminaire

D'après les données, il semble intuitif de conjecturer que le temps d'attente moyen d'un usager à l'arrêt de bus sera de 7.5 minutes. En effet, on sait qu'en moyenne les intervalles X_n ont pour longueur 15 minutes. La logique voudrait, sachant qu'un utilisateur arrivera à un horaire t tiré au sort de manière uniforme, que la moyenne du temps d'attente soit moitié moindre. Si l'on ne se représente que des intervalles X_n de durée moyenne, alors la moyenne du temps d'attente dans un tel intervalle sera clairement de la moitié de cette durée.

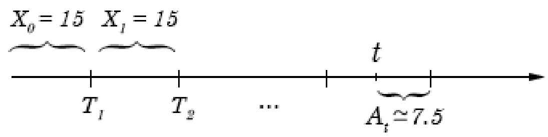

FIGURE 2.1 – Notre première conjecture

2.3 Simulation de A_t

2.3.1 Algorithme de simulation

On rappelle que A_t est la variable aléatoire représentant le temps à attendre avant l'arrivée d'un bus. Simuler A_t, après avoir simulé un horaire $t > 0$ de manière uniforme, c'est donc :

1. Simuler le processus de Poisson en générant et sommant les X_n qui suivent une loi exponentielle de paramètre λ.
2. S'arrêter quand l'intervalle dans lequel se trouve t est atteint : il s'agit de $[T_{N_t}; T_{N_t+1}[$.
3. En déduire $A_t = T_{N_t+1} - t$.

Nous allons formaliser cet algorithme en pseudo-language. On suppose que l'on connaît une fonction $simulerUneLoiUniforme(a, b)$ qui nous permet de simuler $U \hookrightarrow \mathcal{U}_{[a;b]}$. Lors de l'implémentation en langage C de cet algorithme, nous utiliserons la fonction $rand()$ pour en déduire $U = a + (b - a)rand()/((double)RAND\_MAX))$.

fonction simulerUneLoiExponentielle (lambda : **Reel) : Reel**
debut
    retourner -ln(simulerUneLoiUniforme(0,1))/lambda
fin

FIGURE 2.2 – L'algorithme de simulation d'une loi exponentielle

fonction simulerAt (t : **Reel**, lambda : **Reel) : Reel**
    Déclaration somme : **Reel**
debut
    somme ← 0
    tant que (somme < t) **faire**
        somme ← somme + simulerUneLoiExponentielle(lambda)
    fintantque
    retourner (somme - t)
fin

FIGURE 2.3 – L'algorithme de simulation de A_t

Il ne nous restera qu'à faire la moyenne d'un certain nombre de simulations de A_t pour avoir un ordre d'idée de $\mathbb{E}(A_t)$, pour ainsi savoir si notre conjecture est correcte.

fonction simulerEsperanceAt (t : **Reel**, lambda : **Reel**, nbEssais : **NaturelNonNul) : Reel**
  Déclaration somme : **Reel**, i : **NaturelNonNul**
debut
  somme ← 0
  pour i ←1 **à** nbEssais **faire**
    somme ← somme + simulerAt(t, lambda)
  finpour
  retourner somme/nbEssais
fin

FIGURE 2.4 – L'algorithme de calcul de l'espérance de A_t

2.3.2 Implémentation de l'algorithme en langage C

On se propose de réaliser un petit programme en C pour estimer l'espérance de A_t. En plus d'être un langage portable, important pour comprendre le fonctionnement d'un ordinateur (on entend par là proche de la machine, ou « langage de bas niveau »), le C est aussi modulaire. C'est pourquoi nous avons ici séparé notre code source en différents modules, ce qui nous permet deux choses :

1. Une meilleure répartition du travail, le programme étant scindé en « boîtes noires » pour que chaque membre du binôme puisse concevoir sa partie sans se soucier du travail de l'autre
2. Une séparation plus claire et plus efficace pour une meilleure réutilisabilité (ce ne sera pas notre dernier projet en C, autant prévoir pour les prochains).

Ainsi, on peut distinguer différents fichiers où sont regroupées fonctions et procédures spécifiques à certaines tâches :

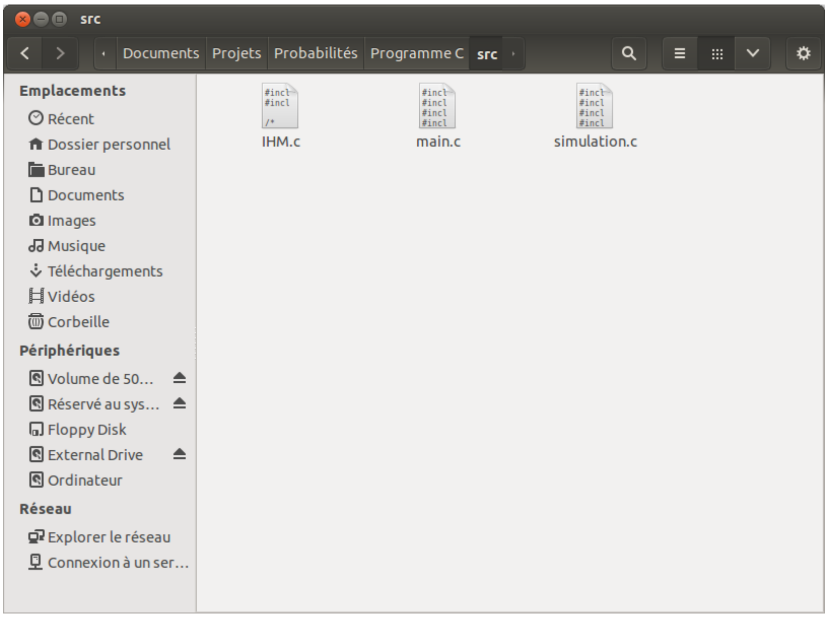

FIGURE 2.5 – Les fichiers source utilisés par notre programme

— Le *main*
— Le module *IHM* (Interface Homme Machine) où sont regroupées des fonctions de saisie et d'affichage sur l'entrée et la sortie standard
— Le module *simulation*, qui, comme son nom l'indique, contient les fonctions de simulation de nos différentes variables aléatoires
— Les trois fichiers d'en-têtes (.h) situés dans le dossier *include* où sont énoncés les constantes et prototypes des fonctions
— Et le fameux makefile pour nous simplifier la compilation, qui nous génère un exécutable dans le dossier *bin*.

Notre programme demande à l'utilisateur la valeur de λ et le nombre de simulations de A_t qu'il souhaite effectuer pour déterminer l'espérance. Nous allons ici l'essayer avec notre valeur de lambda $\lambda = \frac{1}{15} \approx 0.06666666666$. On commence en calculant l'espérance de A_t sur 10000 simulations, ce qui devrait suffir amplement :

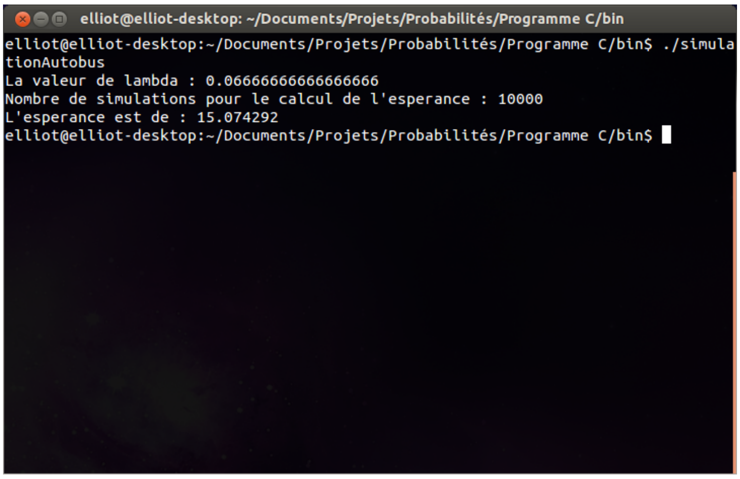

FIGURE 2.6 – Un premier calcul de $\mathbb{E}[A_t]$

C'est une valeur très étrange et très éloignée de celle estimée... Nous allons donc essayer de réitérer le calcul...

2.4 Approximation et conjecture de $\mathbb{E}[A_t]$

Au vu de la première valeur renvoyée par le programme, on peut penser que le hasard nous a joué un tour, on recommence donc l'opération mais cette fois-ci sur un million de simulations :

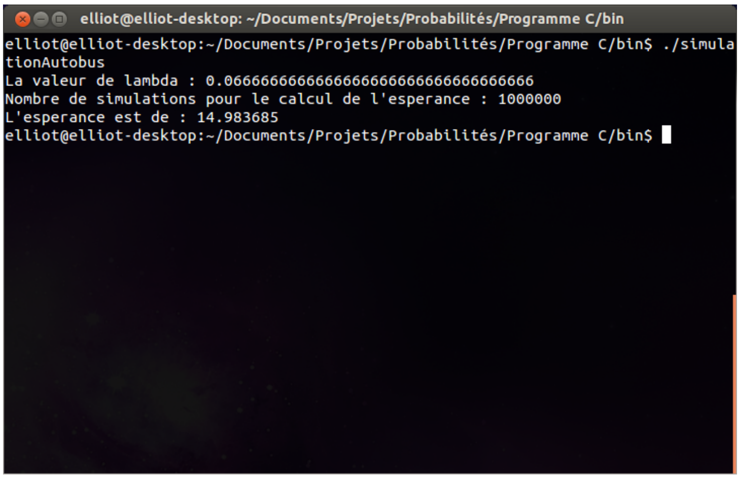

FIGURE 2.7 – Un deuxième calcul de $\mathbb{E}[A_t]$

Bigre ! Il semble que $\mathbb{E}[A_t] \approx 15$, c'est très surprenant car il s'agit de la moyenne de l'écart temporel entre deux bus.

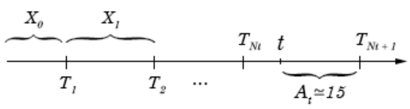

FIGURE 2.8 – Représentation de $\mathbb{E}[A_t]$

Cela semble incroyable, on aurait donc :

$$\mathbb{E}[A_t] = \frac{1}{\lambda}$$

C'est très paradoxal dans le sens où un usager arrivant à un horaire t attendra aussi longtemps (voir plus longtemps) son bus que s'il venait d'en rater un... La compagnie gérant les autobus peut donc se féliciter d'avoir un écart moyen de 15 minutes entre chaque passage d'autobus, mais cela ne joue pas forcément en la faveur de leur clientèle ! On peut assumer que ce résultat surprenant est dû à la distribution exponentielle des bus. Notre conjecture aurait-elle pu être la bonne si l'on avait supposé que l'écart moyen entre deux bus suivait une loi uniforme ? En changeant la simulation de A_t par une loi uniforme sur un intervalle $[0; 30]$, il en vient que :

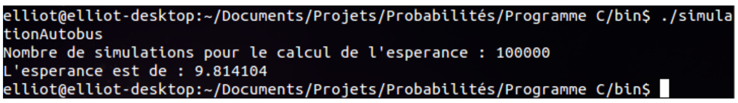

FIGURE 2.9 – Un calcul de $\mathbb{E}[A_t]$ avec des accroissements différents

Non plus ! Décidément, le temps d'attente est toujours plus grand que l'espérance des X_n. On peut donc imaginer que cela vient du fait que lorsque l'on simule un horaire, il a beaucoup plus de chance de tomber dans un grand intervalle de temps que dans un petit.

Calcul de l'espérance de A_t

Dans cette partie, nous commencerons par calculer la loi de A_t pour pouvoir ensuite étudier son espérance. On essayera enfin de lever complètement le paradoxe concernant l'espérance du temps d'attente en expliquant la raison pour laquelle l'intuition feint un résultat erroné.

3.1 Détermination de la loi de A_t

On sait que A_t est le temps d'attente avant le prochain bus. Autrement dit, si l'on a raté le $n$ième bus, cela correspond au temps à attendre avant d'attraper le $(n+1)$ième bus :

$$\begin{aligned}\forall \omega \in \Omega,\ \forall x \in \mathbb{R}_+,\ \omega \in \{A_t \leq x\} &\Leftrightarrow \exists n \in \mathbb{N} \ / \ \omega \in \{T_{n+1} > t \geq T_n\} \ et \ T_{n+1}(\omega) - t \leq x \\ &\Leftrightarrow \exists n \in \mathbb{N} \ / \ \omega \in \{N_t = n\} \cap \{T_{n+1} - t \leq x\} \\ &\Leftrightarrow \omega \in \bigcup_{n \geq 0} \{N_t = n\} \cap \{T_{n+1} - t \leq x\} \\ &\Leftrightarrow \omega \in \{T_{N_t+1} - t \leq x\}\end{aligned}$$

Ainsi :

$$\boxed{A_t = T_{N_t+1} - t}$$

$$\begin{array}{rcccc} A_t & : & \Omega & \to & \mathbb{R}_+ \\ & & \omega & \mapsto & T_{N_t+1}(\omega) - t \end{array}$$

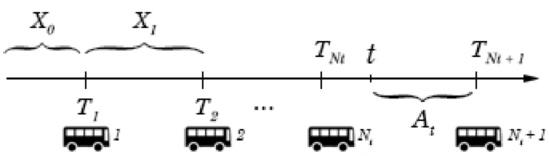

FIGURE 3.1 – Un schéma représentant A_t

On cherche la *f.d.r* de A_t, c'est-à-dire : $\forall x \in \mathbb{R},\ F_{A_t}(x) = \mathbb{P}(A_t \leq x) = 1 - \mathbb{P}(A_t > x)$. Nous allons ici plutôt nous intéresser à la fonction de survie de A_t, c'est-à-dire à $\mathbb{P}(A_t > x)$. On commence par remarquer que les $\{N_t = n\}$ forment une partition de Ω, en effet :

1. $\forall (n,m) \in \mathbb{N}^2,\ (n \neq m \Rightarrow \{N_t = n\} \cap \{N_t = m\} = \varnothing)$
2. $\Omega = \bigcup_{n \geq 0} \{N_t = n\}$

Ainsi,

$$\forall x \in \mathbb{R}_+,\ \mathbb{P}(A_t > x) = \mathbb{P}\left(\bigcup_{n \geq 0} \{A_t > x\} \cap \{N_t = n\}\right)$$

Soit $n \in \mathbb{N}$ fixé. Calculons d'abord la probabilité de l'événement $\{A_t > x\} \cap \{N_t = n\}$:

$$\begin{aligned}\mathbb{P}(\{T_{N_t+1} - t > x\} \cap \{N_t = n\}) &= \mathbb{P}(\{T_{n+1} > t + x\} \cap \{N_t = n\}) \\ &= \mathbb{P}(\{T_{n+1} - T_n > t + x - T_n\} \cap \{N_t = n\}) \\ &= \mathbb{P}(\{X_n > t + x - T_n\} \cap \{N_t = n\})\end{aligned}$$

Or,

$$\begin{aligned}\omega \in \{X_n > t + x - T_n\} \cap \{N_t = n\} &\Leftrightarrow X_n(\omega) > t + x - T_n(\omega) \ et \ T_{n+1}(\omega) > t \geq T_n(\omega) \\ &\Leftrightarrow X_n(\omega) > x \ et \ N_t(\omega) = n\end{aligned}$$

Donc, comme les variables aléatoires N_t et X_n sont clairement indépendantes (le nombre de bus passés à l'instant t ne dépend pas de l'écart temporel avec le prochain bus, ou encore la probabilité d'avoir un écart temporel supérieur à x sachant que n bus sont passés est la même que la probabilité d'avoir un écart temporel supérieur à x) :

$$\begin{aligned}\mathbb{P}(\{T_{N_t+1} - t > x\} \cap \{N_t = n\}) &= \mathbb{P}(\{X_n > x\} \cap \{N_t = n\}) \\ &= \mathbb{P}(X_n > x)\mathbb{P}(N_t = n) \\ &= \left(1 - \left(1 - e^{-\lambda x}\right)\right) e^{-\lambda t} \frac{(\lambda t)^n}{n!} \\ &= e^{-\lambda(t+x)} \frac{(\lambda t)^n}{n!}\end{aligned}$$

Et donc, finalement :

$$\begin{aligned}\mathbb{P}(A_t > x) &= \mathbb{P}\left(\bigcup_{n \geq 0} \{A_t > x\} \cap \{N_t = n\}\right) \\ &= \sum_{n \geq 0} \mathbb{P}(\{A_t > x\} \cap \{N_t = n\}) \\ &= \sum_{n \geq 0} e^{-\lambda(t+x)} \frac{(\lambda t)^n}{n!} \\ &= e^{-\lambda(t+x)} \sum_{n \geq 0} \frac{(\lambda t)^n}{n!} \\ &= e^{-\lambda(t+x)} e^{\lambda t} \\ &= e^{-\lambda x}\end{aligned}$$

D'où :

$$\boxed{\forall x \in \mathbb{R},\ F_{A_t}(x) = \mathbb{P}(A_t \leq x) = (1 - e^{-\lambda x})\mathbb{1}_{\mathbb{R}_+}(x)}$$

On remarque que $A_t \hookrightarrow \xi(\lambda)$ (car une loi est caractérisée par sa $f.d.r$). La loi de A_t aurait pu être calculée un peu plus rapidement [1] en utilisant une variable aléatoire de comptage sur un intervalle quelconque.

3.2 Calcul de $\mathbb{E}[A_t]$

$A_t \hookrightarrow \xi(\lambda)$ ainsi, d'après le calcul effectué précédemment sur l'espérance d'une variable aléatoire exponentielle, on a :

$$\boxed{\mathbb{E}[A_t] = \frac{1}{\lambda}}$$

Dans notre cas, cela nous donne :

$$\boxed{\mathbb{E}[A_t] = 15}$$

Le temps d'attente moyen d'un usager à son arrêt est donc de 15 minutes.

1. Ce calcul est réalisé dans l'annexe C

3.3 Comparaison avec les résultats précédents

On retrouve bien le résultat de notre deuxième conjecture, qui était en désaccord complet avec notre conjecture préliminaire. On avait émis comme hypothèse que cela était dû au fait que t se retrouverait souvent dans un intervalle plus grand. Autrement dit, que la longueur moyenne de l'intervalle contenant t serait plus grande que la longueur moyenne d'un intervalle.

Pour vraiment lever le voile sur ce résultat paradoxal on se propose de calculer, justement, l'espérance de cet intervalle. Pour cela, on pose B_t la variable aléatoire qui définit, lorsqu'un utilisateur arrive à une date $t > 0$ fixée, le temps déjà écoulé depuis le passage du bus qu'il a raté. Autrement dit,

$$\begin{aligned}
\forall \omega \in \Omega,\ \forall x \in \mathbb{R}_+,\ \omega \in \{B_t \leq x\} &\Leftrightarrow \exists n \in \mathbb{N} \ / \ \omega \in \{T_{n+1} > t \geq T_n\} \ et \ t - T_n(\omega) \leq x \\
&\Leftrightarrow \exists n \in \mathbb{N} \ / \ \omega \in \{N_t = n\} \cap \{t - T_n \leq x\} \\
&\Leftrightarrow \omega \in \bigcup_{n \geq 0} \{N_t = n\} \cap \{t - T_n \leq x\} \\
&\Leftrightarrow \omega \in \{t - T_{N_t} \leq x\}
\end{aligned}$$

Ainsi :

$$\boxed{B_t = t - T_{N_t}}$$

Et on remarque, justement, que :

$$\boxed{A_t + B_t = X_{N_t}}$$

La loi de $A_t + B_t$ est exactement la loi de l'intervalle contenant t. Par linéarité de l'espérance, il ne nous reste donc qu'à trouver la loi et l'espérance de B_t. On calcule, là encore, la fonction de survie de B_t, à $x \in \mathbb{R}_+$ fixé :

$$\begin{aligned}
\mathbb{P}(B_t > x) &= \mathbb{P}\left(\bigcup_{n \geq 0} \{B_t > x\} \cap \{N_t = n\}\right) \\
&= \sum_{n \geq 0} \mathbb{P}(\{B_t > x\} \cap \{N_t = n\}) \\
&= \sum_{n \geq 0} \mathbb{P}(\{t - x > T_n\} \cap \{N_t = n\})
\end{aligned}$$

On distingue trois cas :
- Le cas où $x < t$.
- Le cas où $x > t$, ce qui implique que l'événement $\{B_t > x\}$ est équivalent à $\{0 > T_{N_t}\}$, ce qui est évidemment de probabilité nulle. Donc, pour $x > t$, $\mathbb{P}(\{B_t \leq x\}) = 1$.
- Le cas où $x = t$.

Pour $x < t$:

$$\begin{aligned}
\mathbb{P}(B_t > x) &= \sum_{n\geq 0} \mathbb{P}(\{t - x > T_n\} \cap \{N_t = n\}) \\
&= \sum_{n\geq 0} \mathbb{P}(\{t - x > T_n\} \cap \{T_n \leq t < T_{n+1}\}) \\
&= \sum_{n\geq 0} \mathbb{P}(\{x - t + T_{n+1} < T_{n+1} - T_n\} \cap \{N_t = n\}) \\
&= \sum_{n\geq 0} \mathbb{P}(\{x < X_n\} \cap \{N_t = n\}) \\
&= \sum_{n\geq 0} \mathbb{P}(\{x < X_n\})\mathbb{P}(\{N_t = n\}) \\
&= \sum_{n\geq 0} e^{-\lambda x} e^{-\lambda t} \frac{(\lambda t)^n}{n!} \\
&= e^{-\lambda x}
\end{aligned}$$

Calculons $\mathbb{P}(B_t \leq t)$:

$$\begin{aligned}
\mathbb{P}(B_t \leq t) &= \mathbb{P}(\{T_{N_t} \leq 0\}) \\
&= \mathbb{P}(\{T_{N_t} = 0\}) \\
&= \mathbb{P}(\{N_t = 0\}) \\
&= e^{-\lambda t}
\end{aligned}$$

D'où :

$$\boxed{\forall x \in \mathbb{R},\ \mathbb{P}(B_t \leq x) = (1 - e^{-\lambda x})\mathbb{1}_{0\leq x<t}(x) + \mathbb{1}_{x>t}(x) + e^{-\lambda t}\mathbb{1}_{x=t}(x)}$$

C'est une loi exponentielle tronquée en t.

On en déduit la densité de B_t :

$$\begin{aligned}
\forall x \in \mathbb{R} \setminus \{t\}, f_{B_t}(x) &= F'_{B_t}(x) \\
&= \lambda e^{-\lambda x}\mathbb{1}_{0\leq x<t}(x)
\end{aligned}$$

Au point $x = t$, la fonction de répartition n'est pas continue et donc la densité n'existe pas. On dit que B_t admet un atome de probabilité en t : $\mathbb{P}(B_t \leq t) = \mathbb{P}(B_t = t) \neq 0$.

Calculons $\mathbb{E}[B_t]$, par le théorème du transfert :

$$\begin{aligned}
\mathbb{E}[B_t] &= \int_{-\infty}^{\infty} x f_{B_t}(x)dx \\
&= \int_0^t \lambda x e^{-\lambda x}dx + tP(B_t = t) \\
&= \lambda\left(\left[-\frac{x}{\lambda}e^{-\lambda x}\right]_0^t - \int_0^t -\frac{1}{\lambda}e^{-\lambda x}dx\right) + te^{-\lambda t} \\
&= \lambda\left(-\frac{t}{\lambda}e^{-\lambda t} + \left[-\frac{1}{\lambda^2}e^{-\lambda x}\right]_0^t\right) + te^{\lambda t} \\
&= \frac{1}{\lambda}(1 - e^{-\lambda t})
\end{aligned}$$

Finalement,

$$\boxed{\mathbb{E}[B_t] = \frac{1}{\lambda}(1 - e^{-\lambda t})}$$

D'où :

$$\boxed{\mathbb{E}[X_{N_t}] = \mathbb{E}[A_t] + \mathbb{E}[B_t] = \frac{1}{\lambda}(2 - e^{-\lambda t})}$$

Ainsi, on voit que l'intervalle contenant t est plus grand en moyenne qu'un intervalle quelconque ! Et même, quand $t \to \infty$, il devient de longueur moyenne double !

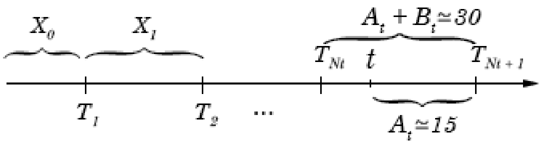

FIGURE 3.2 – L'espérance de la longueur de l'intervalle contenant t

Donc un usager arrive très souvent « au mauvais moment », surtout s'il arrive tard. Cela nous fait un peu penser à la loi de Murphy : « Tout ce qui peut mal tourner, tournera mal ».

Conclusion

Ce projet, d'une durée relativement limitée, nous aura néanmoins enseigné de nombreuses choses :

— Une maîtrise plus poussée des probabilités, une meilleure compréhension du cours.
— De nouvelles connaissances et une introduction au processus de Poisson que l'on verra en GM4.
— Une opportunité pour nous perfectionner en langage C.
— La compréhension d'un paradoxe surprenant.
— Une vision plus globale de la matière Probabilités-Statistiques (GM3) que nous continuerons d'étudier toute l'année.

Le calcul probabiliste est une branche qui nous tient à coeur dans le sens où l'on peut souvent appliquer la théorie mathématique à des situations bien concrètes. Les outils de calcul des probabilités mis à notre disposition nous ont permis ici de complètement lever le voile sur une situation très paradoxale.

On a pu notamment mettre en évidence que le temps d'attente d'un passager à son arrêt de bus dépassait souvent les limites de sa patience, expliquant ainsi les vagues de colère fréquentes des usagers des réseaux de transports en commun.

La variable aléatoire de comptage dans un intervalle $N_{]a;b]}$

Nous allons ici généraliser le résultat trouvé pour $N_t = N_{[0;t]}$ dans la première partie de ce projet :

$$\forall n \in \mathbb{N},\ \mathbb{P}(N_t = n) = e^{-\lambda t}\frac{(\lambda t)^n}{n!}$$

On admet ici que le processus de Poisson est stationnaire, c'est-à-dire que $\forall (s < t) \in \mathbb{R}_+^2$:

$$N_{t+s} - N_s \hookrightarrow N_t$$

On remarque ensuite que :

$$\begin{aligned} N_b = N_a + N_{]a;b]} &\Leftrightarrow N_{]a;b]} = N_{[0;b]} - N_{[0;a]} \\ &\Leftrightarrow N_{]a;b]} = N_{[0;a+(b-a)]} - N_{[0;a]} \end{aligned}$$

Comme $(b-a) \in \mathbb{R}^+$, on a :

$$\boxed{N_{]a;b]} \hookrightarrow N_{(b-a)}}$$

Et donc :

$$\boxed{\forall n \in \mathbb{N},\ \mathbb{P}(N_{]a;b]} = n) = e^{-\lambda(b-a)}\frac{(\lambda(b-a))^n}{n!}}$$

Une autre manière de calculer I_n

On rappelle que l'on voulait calculer, pour tout $n \in \mathbb{N}^*$ et $x \in \mathbb{R}_+$ fixé,

$$I_n(x) = \int_0^x t^{n-1} e^{-\lambda t} dt$$

Une façon astucieuse de résoudre ce problème est de poser, pour un polynôme P de degré n fixé :

$$\forall x \in \mathbb{R},\ D(x) = -\left[\sum_{k=0}^{n} P^{(k)}(x)\right] e^{-x}$$

Et alors, on peut remarquer que :

$$\begin{aligned}
D'(x) &= -\left[\sum_{k=0}^{n-1} P^{(k+1)}(x) - \sum_{k=0}^{n} P^{(k)}(x)\right] e^{-x} \\
&= -\left[\sum_{k=1}^{n} P^{(k)}(x) - \sum_{k=0}^{n} P^{(k)}(x)\right] e^{-x} \\
&= P(x) e^{-x}
\end{aligned}$$

On peut déduire de l'égalité précédente que :

$$\begin{aligned}
\int_0^x P(t) e^{-t} dt &= \int_0^x D'(t) dt \\
&= D(x) - D(0) \\
&= -\left[\sum_{k=0}^{n} P^{(k)}(x)\right] e^{-x} + \sum_{k=0}^{n} P^{(k)}(0)
\end{aligned}$$

On se donne le polynôme P tel que $\forall t \in \mathbb{R},\ P(t) = t^{n-1}$ de degré $n-1$. Alors, en appliquant la formule ci-dessus :

$$\begin{aligned}
\int_0^x t^{n-1} e^{-t} dt &= -\left[x^{n-1} + (n-1)x^{n-2} + ... + (n-1)...2x^1 + (n-1)...1\right] e^{-x} + 0 + (n-1)...1 \\
&= -\left[\sum_{k=0}^{n-1} \frac{(n-1)!}{(n-1-k)!} x^{n-1-k}\right] e^{-x} + (n-1)! \\
&= -\left[\sum_{k=0}^{n-1} \frac{(n-1)!}{k!} x^{k}\right] e^{-x} + (n-1)!
\end{aligned}$$

Par un changement de variable $u = \lambda t$ dans I_n, on a :

$$\begin{aligned}
I_n(x) &= \int_0^x t^{n-1} e^{-\lambda t} dt \\
&= \int_0^{\lambda x} (\frac{u}{\lambda})^{n-1} e^{-u} \frac{1}{\lambda} du \\
&= \frac{1}{\lambda^n} \int_0^{\lambda x} u^{n-1} e^{-u} du \\
&= \frac{1}{\lambda^n} \left(- \left[\sum_{k=0}^{n-1} \frac{(n-1)!}{k!} (\lambda x)^k \right] e^{-\lambda x} + (n-1)! \right) \\
&= - \left[\sum_{k=0}^{n-1} \frac{(n-1)!}{k!} \frac{(\lambda x)^k}{\lambda^n} \right] e^{-\lambda x} + \frac{(n-1)!}{\lambda^n}
\end{aligned}$$

Ainsi on retrouve bien le résultat que l'on avait démontré par récurrence :

$$\boxed{\forall n \in \mathbb{N}^*,\ \forall x \in \mathbb{R},\ I_n(x) = \left(\frac{(n-1)!}{\lambda^n} - \left[\sum_{k=0}^{n-1} \frac{(n-1)!}{k!} \frac{(\lambda x)^k}{\lambda^n} \right] e^{-\lambda x} \right) \mathbb{1}_{\mathbb{R}_+}(x)}$$

Une autre façon de déterminer la loi de A_t et de B_t

On cherche ici la fonction de répartition de A_t et de B_t. La démonstration est presque la même, sauf qu'ici on utilise la variable aléatoire de comptage sur un intervalle, ce qui nous donne la solution un tout petit peu plus rapidement. Encore une fois, on se ramène à la fonction de survie $\mathbb{P}(A_t > x)$. Soit $x \in \mathbb{R}^+$:

$$\begin{aligned}
\mathbb{P}(A_t > x) &= \mathbb{P}\left(\bigcup_{n\geq 0}\{A_t > x\} \cap \{N_t = n\}\right) \\
&= \sum_{n\geq 0} \mathbb{P}(\{T_{n+1} > t + x \geq t \geq T_n\}) \\
&= \sum_{n\geq 0} \mathbb{P}(\{N_{]t;t+x]} = 0\} \cap \{N_t = n\}) \\
&= \sum_{n\geq 0} \mathbb{P}(\{N_{]t;t+x]} = 0\})\mathbb{P}(\{N_t = n\})
\end{aligned}$$

Il y a indépendance entre $N_{]t;t+x]}$ et N_t car le nombre de passages d'autobus entre t et $(t + x)$ ne dépend pas du nombre de bus passés jusqu'à la date t.

$$\begin{aligned}
\mathbb{P}(A_t > x) &= \sum_{n\geq 0} \mathbb{P}(\{N_{]t;t+x]} = 0\})\mathbb{P}(\{N_t = n\}) \\
&= \sum_{n\geq 0} e^{-\lambda(t+x-t)} e^{-\lambda t} \frac{(\lambda t)^n}{n!} \\
&= e^{-\lambda x}
\end{aligned}$$

D'où, pour tout $x \in \mathbb{R}$:

$$\boxed{\mathbb{P}(A_t \leq x) = (1 - e^{-\lambda x})\mathbb{1}_{\mathbb{R}_+}(x)}$$

On retrouve bien la loi de A_t trouvée en troisième partie de ce projet.

Pour ce qui est de B_t, on distingue encore les trois cas : $x < t$, $x > t$ et $x = t$.

Soit $x < t$:

$$\begin{aligned}
\mathbb{P}(B_t > x) &= \mathbb{P}\left(\bigcup_{n\geq 0}\{B_t > x\} \cap \{N_t = n\}\right) \\
&= \sum_{n\geq 0} \mathbb{P}(\{t - x > T_n\} \cap \{T_{n+1} > t \geq T_n\}) \\
&= \sum_{n\geq 0} \mathbb{P}(\{T_{n+1} > t > t - x \geq T_n\}) \\
&= \sum_{n\geq 0} \mathbb{P}(\{N_{]t-x;t]} = 0\})\mathbb{P}(\{N_{t-x} = n\}) \\
&= \sum_{n\geq 0} e^{-\lambda(t-(t-x))} e^{-\lambda(t-x)} \frac{(\lambda(t-x))^n}{n!} \\
&= e^{-\lambda x}
\end{aligned}$$

Pour $x > t$ et $x = t$, on a toujours respectivement $\mathbb{P}(\{B_t \leq x\}) = 1$ et $\mathbb{P}(\{B_t \leq t\}) = e^{-\lambda t}$. D'où, pour tout $x \in \mathbb{R}$:

$$\boxed{\mathbb{P}(B_t \leq x) = (1 - e^{-\lambda x})\mathbb{1}_{0\leq x<t}(x) + \mathbb{1}_{x>t}(x) + e^{-\lambda t}\mathbb{1}_{x=t}(x)}$$

On retrouve bien la loi exponentielle tronquée en t précédemment calculée.

Programme de simulation

D.1 Les headers

D.1.1 IHM

Fichier : IHM.h

```
/*
    Module : IHM
    Utilité : Intéragir avec l'utilisateur
*/

double recupererLambda();
int recupererNbEssais();
void afficherEsperance(double esp);
```

D.1.2 Simulation

Fichier : simulation.h

```
/*
    Module : simulation
    Utilité : Simuler des lois de probabilités
*/

double simulerUneLoiUniforme(double a, double b);
double simulerUneLoiExp(double lambda);
double simulerUnHoraire();
double simulerAt(double t, double lambda);
double simulerEsperanceAt(double t, double lambda, int nbEssais);
```

D.2 Programme principal

Fichier : main.c

```
#include <stdlib.h>
#include <time.h>
#include "IHM.h"
#include "simulation.h"

/*
    Programme principal
*/

int main(int argc, char* argv[]){
    /* Saisie utilisateur */
    double lambda = recupererLambda();
    int nbEssais = recupererNbEssais();
```

```

    /* Initialisation de la graine */
    srand(time(NULL));

    /* Calcul et affichage de l'espérance */
    afficherEsperance(simulerEsperanceAt(simulerUnHoraire(), lambda, nbEssais));

    return EXIT_SUCCESS;
}
```

D.3 Le module IHM

Fichier : IHM.c

```
#include <stdio.h>
#include "IHM.h"

/*
    Récuperer lambda
    Sortie : Un réel positif lambda saisi par l'utilisateur
*/
double recupererLambda(){
    double lambda;
    printf("La valeur de lambda : ");
    scanf("%lf", &lambda);
    return lambda;
}

/*
    Récuperer le nombre d'essais
    Sortie : Un entier positif nbEssais saisi par l'utilisateur
*/
int recupererNbEssais(){
    int nbEssais;
    printf("Nombre de simulations pour le calcul de l'esperance : ");
    scanf("%d", &nbEssais);
    return nbEssais;
}

/*
    Afficher l'espérance
    Entrée : Un réel esp
*/
void afficherEsperance(double esp){
    printf("L'esperance est de : %f \n", esp);
}
```

D.4 Le module simulation

Fichier : simulation.c

```
#include <stdlib.h>
#include <math.h>
#include <time.h>
```

```c
#include "simulation.h"

/*
    Constante : temps maximal
*/
double const T_MAX = 1000;


/*
    Simuler une loi uniforme
    Entrée : l'intervalle [a,b]
    Sortie : Un réel au hasard dans cet intervalle
*/
double simulerUneLoiUniforme(double a, double b){
    return (a + (b-a)*(rand()/((double)RAND_MAX)));
}

/*
    Simuler un horaire
    Sortie : Un réel au hasard dans l'intervalle [0,T_MAX]
*/
double simulerUnHoraire(){
    return (simulerUneLoiUniforme(0,T_MAX));
}

/*
    Simuler une loi exponentielle
    Entrée : Un réel lambda
    Sortie : Un réel au hasard suivant la loi exponentielle de paramètre lambda
*/
double simulerUneLoiExp(double lambda){
    return (-log(simulerUneLoiUniforme(0,1))/lambda);
}

/*
    Simuler la variable aléatoire At
    Entrée : Un horaire réel positif t
             Un réel lambda
    Sortie : Un réel au hasard en suivant la loi de At
*/


double simulerAt(double t, double lambda){
    double sum = 0;
    while (t > sum)
        sum += simulerUneLoiExp(lambda);
    return (sum - t);
}

/*
    Simuler l'espérance de la variable aléatoire At
    Entrée : Un horaire réel positif t
             Un réel lambda
             Un nombre de simulations de At
    Sortie : Un réel calculant l'espérance de At
*/
double simulerEsperanceAt(double t, double lambda, int nbEssais){
    int i = 0;
    double sum = 0;
    for (i = 0 ; i < nbEssais ; i++)
        sum += simulerAt(t, lambda);
    return (sum/nbEssais);
```

```
}
```

D.5 Le makefile

Fichier : Makefile

```
# Variables dossiers
SRCDIR=src
BINDIR=bin
INCLUDEDIR=include

# Compilateur
CC=gcc

# Options de compilation
CFLAGS=-Wall -pedantic -I$(INCLUDEDIR)

# Variables sources et objets
SRC=$(SRCDIR)/main.c $(SRCDIR)/IHM.c $(SRCDIR)/simulation.c
OBJ= $(SRC:.c=.o)

# Nom du programme
EXEC=simulationAutobus


# Commande par défaut
all: $(BINDIR)/$(EXEC)

# Programme principal
$(BINDIR)/$(EXEC): $(OBJ)
	@$(CC) -o $@ $^ -lm

# Éditions de liens
$(SRCDIR)/%.o: $(SRCDIR)/%.c
	@$(CC) -o $@ -c $< $(CFLAGS)

# Nettoyage
clean:
	@rm $(SRCDIR)/*.o

mrproper: clean
	@rm $(BINDIR)/*
```

Bibliographie

[1] *Mohamed El Machkouri*, *Cours de Probabilités/Statistiques*, Institut National des Sciences Appliquées de Rouen.

[2] *E. Lecarbier, S. Robin*, *Un bref aperçu du processus de Poisson*, AgroParisTech, 2007.

[3] *F. Testard*, *Modèles exponentiels*, Université de La Rochelle, 2010.

[4] `http://www.universalis.fr/encyclopedie/simeon-denis-poisson/` (Valide à la date du 20/11/2013) Biographie de Siméon Denis Poisson.

[5] `http://nestor.coventry.ac.uk/~nhunt/poisson/history.html` (Valide à la date du 20/11/2013) Une bref histoire concernant la distribution de Poisson.

[6] `http://images.math.cnrs.fr/Poisson.html` (Valide à la date du 20/11/2013) Quelques détails sur l'ouvrage de Poisson « Recherches sur la probabilité des jugements en matière criminelle et matière civile ».

[7] `http://www.math.nyu.edu/faculty/varadhan/spring06/spring06.1.pdf` (Valide à la date du 20/11/2013) Un pdf bien détaillé sur le processus de Poisson.

[8] `http://www.math.uah.edu/stat/poisson/` (Valide à la date du 20/11/2013) Un cours sur le processus de Poisson.

[9] `http://www.netlab.tkk.fi/opetus/s383143/kalvot/E_poisson.pdf` (Valide à la date du 20/11/2013) Un pdf avec des explications du processus de Poisson et ses propriétés.