Fonctions génératrices, Fonctions caractéristiques, Convolution

1 Fonctions génératrices

Si X est une variable aléatoire à valeurs dans $\mathbb{N}$, la fonction génératrice de X est définie comme

$$G_X(s) = \mathbb{E}(s^X) = \sum_{k=0}^{\infty} p_k s^k$$

où $p_k = \mathbb{P}(X = k)$ et où on adopte la convention $0^0 = 1$.

Proposition 1.1. *La fonction génératrice vérifie les propriétés suivantes*
(*i*) *Son rayon de convergence est supérieur ou égal à* 1 *et elle est* C^∞ *sur* $(-1; 1)$.
(*ii*) *Elle caractérise la loi de* X *(i.e. deux fonctions génératrices égales correspondantes à des v.a. de même loi).*
(*iii*) $\lim_{s\to 1^-} G'_X(s) = \mathbb{E}(X)$.
(*iv*) *Si* X *et* Y *sont deux variables aléatoires à valeurs entières indépendantes alors* $G_{X+Y} = G_X.G_Y$

On rappelle que si $\mathbb{E}(X) < \infty$, la dérivée de G_X est prolongeable par continuité en 1 et G_X devient dérivable à gauche en 1.

Théorème 1.1. *Si* X_1, X_2, ... *est une suite de v.a.i.i.d. de fonction génératrice commune* G_X *et si* N *est une variable aléatoire à valeurs dans* $\mathbb{N}$ *indépendante des* X_i *et de fonction génératrice* G_N, *alors* $S_N = X_1 + X_2 + \ldots + X_N$ *a pour fonction génératrice*

$$G_{S_N}(s) = G_N(G_X(s)).$$

1) Moyen mémotechnique : pour se souvenir de l'ordre dans lequel les fonctions sont composées, remarquer que si N est déterministe ($\mathbb{P}(N = n) = 1$) alors on doit retrouver la formule donnant la fonction génératrice de n v.a.i.i.d. soit $G_{S_N}(s) = G_X(s)^n$, et l'on a $G_N(s) = s^n$.

2) Démonstration : Par le théorème de Fubini et indépendance de N et $(X_i : i \in \mathbb{N})$,

$$\begin{aligned} G_{S_N}(s) &= \mathbb{E}([\sum_{n\geq 0} \mathbb{1}(N=n)]s^{X_1+\ldots+X_N}) \\ &= \sum_{n\geq 0} \mathbb{E}(\mathbb{1}(N=n)s^{X_1+\ldots+X_n}) \\ &= \sum_{n\geq 0} \mathbb{P}(N=n)\mathbb{E}(s^{X_1+\ldots+X_n}) \\ &= \sum_{n\geq 0} \mathbb{P}(N=n)\mathbb{E}(s^X)^n \\ &= G_N(G_X(s)). \end{aligned}$$

1.1 Exercices

Exercice 1.1. On rappelle que si X est une variable aléatoire à valeurs dans $\mathbb{N}$, la fonction génératrice de X est définie comme

$$G_X(s) = \mathbb{E}(s^X) = \sum_{k=0}^{\infty} p_k s^k$$

où $p_k = \mathbb{P}(X=k)$ et où on adopte la convention $0^0 = 1$.

0) Rayon de convergence de la série ?

1) Montrer que la fonction génératrice de X caractérise complètement sa loi.

2) Montrer que G_X est C^∞ sur $(-1;1)$ et que si X est intégrable alors

$$\lim_{s\to 1^-} G'_X(s) = \mathbb{E}(X)$$

Donner un résultat analogue pour la variance. Montrer de même que $\mathbb{E}(X(X-1)\ldots(X-k+1)) = G^{(k)}(1-)$. *Indication. On utilisera et montrera le théorème d'Abel : Soit $a_n \geq 0$ et $G(s) = \sum_{n\in\mathbb{N}} a_n s^n$. Si $G(s)$ est fini pour $s < 1$, alors $\lim_{s\to 1} G(s) = \sum_{n\in\mathbb{N}} a_n$.*

3) Montrer que si X et Y sont deux variables aléatoires à valeurs entières indépendantes alors la fonction génératrice de $X+Y$ est le produit de G_X et de G_Y. Réciproque ?

4) Calculer les fonctions génératrices des distributions suivantes : v.a. constante $X = c$, variable de Bernoulli de paramètre p, distribution géométrique ($\mathbb{P}(X=k) = p(1-p)^{k-1}$ pour $k \geq 1$), distribution de Poisson, distribution binomiale $\mathcal{B}(n,p)$. Vérifier sans calcul que la somme de deux variables indépendantes $\mathcal{B}(n,p)$ et $\mathcal{B}(m,p)$ est $\mathcal{B}(n+m,p)$.

5) On appelle fonction génératrice jointe de deux v.a. X_1 et X_2 à valeurs dans $\mathbb{N}$ la fonction

$$G_{X_1,X_2}(s_1,\ s_2) = \mathbb{E}(s_1^{X_1} s_2^{X_2}).$$

Montrer que X_1 et X_2 sont indépendantes si et seulement si

$$G_{X_1,\,X_2}(s_1,\ s_2) = G_{X_1}(s_1)G_{X_2}(s_2)$$

pour tous s_1 et s_2.

Solutions : 0) Le rayon de convergence de la série est supérieur ou égal à 1.

1) Remarquer que l'on a $p_k = \frac{1}{k!} G_X^{(k)}(0)$.

2) G_X *est une série entière et donc en tant que telle* C^∞ *sur son disque de convergence. D'autre part, pour tout* $s \in (-1, 1)$*, on a (en utilisant le théorème d'Abel)*

$$G_X'(s) = \sum_{k=1}^{\infty} k p_k s^{k-1} \uparrow_{s \to 1} \sum_{k=1}^{\infty} k p_k = \mathbb{E}(X)$$

De même, on montre que si X *a un moment d'ordre* 2

$$Var(X) = \lim_{s \to 1} \left[G_X^{(2)}(s) + G_X^{(1)}(s) - [G_X^{(1)}(s)]^2 \right]$$

3) $G_{X+Y}(s) = \mathbb{E}(s^X s^Y) = \mathbb{E}(s^X)\mathbb{E}(s^Y) = G_X(s)G_Y(s)$ *par indépendance de* X *et* Y*. La réciproque est fausse (comme elle l'est pour les fonctions caractéristiques, cf Exo 2.5). Ce qui est en revanche vrai, c'est que si la fonction génératrice du couple* (X, Y) *(définie par* $G_{(X,Y)}(s_1, s_2) = \mathbb{E}(s_1^X s_2^Y)$*) s'écrit comme le produit direct de* G_X *et* G_Y *(i.e.* $G_{(X,Y)}(s_1, s_2) = G_X(s_1)G_Y(s_2)$*) alors les deux variables* X *et* Y *sont indépendantes (voir la question 5)).*

4) On obtient successivement $G(s) = s^c$ *(variable constante),* $G(s) = (1-p) + ps$ *(variable de Bernoulli),* $G(s) = \sum_{k=1}^{\infty} s^k p(1-p)^{k-1} = \frac{ps}{1-s(1-p)}$ *(variable géométrique de paramètre* p*),* $G(s) = \sum_{k=0}^{\infty} s^k \frac{\lambda^k}{k!} e^{-\lambda} = e^{\lambda(s-1)}$ *(variable de Poisson),* $G(s) = (q + ps)^n$ *(variable binomiale* $\mathcal{B}(n, p)$*,* $q = 1 - p$*.) On peut voir que la somme d'une* $\mathcal{B}(n, p)$ *et d'une* $\mathcal{B}(m, p)$ *indépendantes est une* $\mathcal{B}(n + m, p)$ *de deux manières : d'une part on peut interpréter la somme comme la somme de* $n+m$ *variables de Bernoulli indépendantes. Mais on peut simplement multiplier les fonctions génératrices en utilisant le 3)*

5) Si X_1 *et* X_2 *sont indépendantes,* $s_1^{X_1}$ *et* $s_2^{X_2}$ *aussi et on a donc* $\mathbb{E}(s_1^{X_1} s_2^{X_2}) = \mathbb{E}(s_1^{X_1})\mathbb{E}(s_2^{X_2})$*. Réciproquement, si cette relation est vraie, on a*

$$\sum_{m,n} \mathbb{P}(X_1 = m; X_2 = n) s_1^m s_2^n = (\sum_m \mathbb{P}(X_1 = m) s_1^m)(\sum_n \mathbb{P}(X_2 = n) s_2^n).$$

Il suffit de développer le deuxième membre et d'identifier terme à terme pour déduire que $\mathbb{P}(X_1 = m; X_2 = n) = \mathbb{P}(X_1 = m)\mathbb{P}(X_2 = n)$.

Exercice 1.2. Une poule pond N oeufs, et N est une distribution de Poisson de paramètre λ. Chaque oeuf éclôt avec probabilité p et les éclosions sont des évènements indépendants. Soit K le nombre de poussins. Calculer la fonction génératrice de K et en déduire que K suit une loi de Poisson de paramètre λp.

Solution : On a $G_X(s) = q+ps$, $G_N(s) = \sum_{n=0}^{\infty} s^n \frac{\lambda^n}{n!} e^{-\lambda} = e^{\lambda(s-1)}$*. Donc* $G_K(s) = G_N(G_X(s)) = e^{\lambda(q+ps-1)} = e^{\lambda p(s-1)}$*.* K *suit donc une loi de Poisson de paramètre* λp*.*

1.2 Processus de branchement : l'extinction des grands noms

Source : Grimmett G., Stirzaker D., *Probability and Random Processes*, 5.4 pp. 171-174 On pourra aussi consulter l'Ouvrard tome II p 529 quand on connait les chaines de Markov et les martingales. Ou Modélisation stochastique et simulations - Bercu Chafai, p125, pour des simulations.

Au dix-neuvième siècle, en 1873 exactement, H.W. Watson répond à une question posée par F. Galton la même année. Il s'agit d'expliquer la tragique et progressive disparition des grands noms de l'aristocratie anglaise. Des noms des grands personnages de Shakespeare, bien peu ont laissé descendance : où sont maintenant les Gloucester, et les Buckingham ? Les processus de Galton-Watson, appelés aussi processus de branchement, modélisent la reproduction et permettent de calculer la probabilité d'extinction. On suppose qu'une population évolue par générations et que Z_n désigne le nombre d'individus de la n-ième génération. Chaque membre de la n-ième génération donne le jour à une famille de membres de la $n+1$-ième génération et la taille de cette famille est une variable aléatoire. On fait les hypothèses suivantes :

(a) Les tailles X_i des familles de tous les individus sont des variables indépendantes

(b) Toutes ces tailles X_i ont même loi et on note G leur fonction génératrice et m leur moyenne.

On supposera dans la suite que $Z_0 = 1$. On appelle $G_n(s) = \mathbb{E}(s^{Z_n})$ la fonction génératrice de Z_n. Notre but est de savoir quand la population s'éteint, i.e. Z_n stationne à zéro :

Théorème 1.2. *Quand $n \to \infty$, $\mathbb{P}(Z_n = 0) \to \eta = \mathbb{P}(\exists n \in \mathbb{N} : Z_n = 0)$. La probabilité d'extinction η peut être calculée comme la plus petite racine positive de l'équation $s = G(s)$.*
De plus, $\eta = 1$ ssi $m \leq 1$ et $\mathbb{P}(X = 1) < 1$.

Interpréter ce résultat, que nous allons maintenant prouver (question 3.c)), et observer un exemple de v.a. convergeant vers 0 sans que son intégrale ne le fasse.

Exercice 1.3. 1) Montrer que $G_n(s) = G(G(\dots(G(s))\dots))$, l'itérée n fois de G.

2)On pose $m_n = \mathbb{E}(Z_n)$ et $\sigma_n^2 = Var(Z_n)$. En dérivant une fois, puis deux fois en $s = 1$ la relation $G_n(s) = G(G_{n-1}(s))$, montrer que $\mathbb{E}(Z_n) = m^n$, $Var(Z_n) = n\sigma^2$ si $m = 1$,

$$Var(Z_n) = \frac{\sigma^2(m^n - 1)m^{n-1}}{m-1} \quad \text{si } \ m \neq 1.$$

3) a) Montrer que $\mathbb{P}(Z_n = 0) = G_n(0)$ est une suite croissante qui converge vers η et que

$$\eta = \mathbb{P}(\exists n \in \mathbb{N} : Z_n = 0) = \inf\{s \in [0,1] : G(s) = s\}.$$

b) Montrer que si $m \leq 1$, alors $\eta < 1$ implique $\mathbb{P}(X = 1) = 1$.
c) Prouver le théorème énoncé plus haut.

4) On suppose que les tailles de familles ont la distribution géométrique $\mathbb{P}(X = k) = qp^k$, avec $q = 1 - p$.

a) Avez-vous une interprétation pour cette hypothèse ? Vérifier que $G(s) = q(1-ps)^{-1}$ et, par récurrence, que

$$G_n(s) = \frac{n-(n-1)s}{n+1-ns} \ \text{ si } \ p = q = \frac{1}{2},$$

$$G_n(s) = \frac{q(p^n - q^n - ps(p^{n-1} - q^{n-1}))}{p^{n+1} - q^{n+1} - ps(p^n - q^n)} \ \text{ si } \ p \neq q.$$

b) En déduire que

$$\mathbb{P}(Z_n = 0) = \frac{n}{n+1} \ \text{ si } \ p = q = \frac{1}{2},$$

$$\mathbb{P}(Z_n = 0) = \frac{q(p^n - q^n)}{p^{n+1} - q^{n+1}} \ \text{ si } \ p \neq q.$$

c) En conclure que, sous les hypothèses précédentes, la probabilité d'extinction η de la population des descendants vérifie

$$\eta = 1 \ \text{ si } \ p \leq q, \ \ \eta = \frac{q}{p} \ \text{ si } \ p > q.$$

Obtenir directement ce résultat à partir du théorème.

Solution

1) C'est très simple, mais la démonstration du Grimmett-Stirzaker est pourtant un peu confuse. On a $Z_n = \sum_{i=1}^{Z_n} X_i$, où X_i est la taille de la famille du i-ème individu de la génération $n-1$. Il s'agit d'une somme d'un nombre aléatoire de v.a.i.i.d. de fonctions génératrices G, ce nombre aléatoire Z_{n-1} ayant une loi G_n. Donc par le théorème 1.1, on a $G_n = G_{n-1} \circ G$ et on conclut parce que $G_1 = G$.

2) On rappelle que si X est une v.a. de fonction génératrice G, alors $\mathbb{E}X = G'(1)$ et $var(X) = G''(1) + G'(1) - G'(1)^2$. En dérivant comme indiqué, on a $G_n'(1) = G'(1)G_{n-1}'(1)$ et donc $m_n = mm_{n-1}$, soit $m_n = m^n$. En dérivant deux fois, on obtient $G_n''(s) = G''(G_{n-1}(s))G_{n-1}'(s)^2 + G'(G_{n-1}(s))G_{n-1}''(s)$, soit $G_n''(1) = G''(1)m^{2(n-1)} + mG_{n-1}''(1)$. On en tire la relation de récurrence $\sigma_n^2 = \sigma^2 m^{2(n-1)} + m\sigma_{n-1}^2$ qui donne bien les résultats annoncés.

3) a) La croissante de la suite $u_n = \mathbb{P}(Z_n = 0)$ vient de la croissance de G et du fait que $u_1 \geq 0 = u_0$. Définir ensuite η comme le premier point fixe, et utiliser la stabilité de $[0, \eta]$ pour conclure par croissance majorée. Enfin l'événement $\{Z_n = 0\}$ croit vers $\{\exists n \in \mathbb{N} : Z_n = 0\}$.

b) Faire un dessin : c'est la convexité de G. Utiliser que si $m \leq 1$, la courbe est au dessus de sa tangente donc au dessus de la première bissetrice. Pour qu'il y ait un point fixe avant 1, il faut donc qu'elle se confonde avec la première bissectrice sur un segment, ce qui par analycité implique $G(s) = s$ et donc $\mathbb{P}(X = 1) = 1$.
Si $m > 1$, alors il existe un point fixe strictement inférieur à 1 par TVI puisque $G(0) \geq 0$ et $G(x) < x$ dans un voisinage de 1 (exclu).

4)a) Simple calcul.
b) On a $\mathbb{P}(Z_n = 0) = G_n(0)$, ce qui donne tout de suite le résultat annoncé en utilisant la question précédente.

Passons à l'interprétation, du point de vue de Galton et Watson : si $p \leq q$, on a $p \leq \frac{1}{2}$ et donc la distribution des tailles décroît très vite. De plus on vérifie facilement que l'espérance du nombre de descendants par individu est $G'(1) = \frac{p}{1-p}$. Donc si $p < \frac{1}{2}$, cette espérance est inférieure à 1. Dans ce cas, il y extinction ! Si $p \geq \frac{1}{2}$, l'espérance du nombre de descendants est supérieure à 1 et tend vers l'infini quand $p \to 1$. Dans ce cas, il y a une probabilité de "survie du nom". Mais la probabilité d'extinction reste importante puisque elle est égale à $\frac{q}{p}$. Il peut y avoir extinction même dans une famille prolifique et les grands noms, donc, peuvent disparaître.

2 Fonctions caractéristiques

Bien que l'on présente souvent la transformée de Fourier de manière autonome en probabilités, l'essentiel de la théorie, et en particulier les théorèmes d'inversion, sont plus complétement décrits grâce à la théorie des distributions. Voir par exemple Bony J.M., *Cours d'Analyse*. Dans ce T.D. nous suivrons essentiellement le Tome II d'Ouvrard (chapitre 12) car la fonction caractéristique y est traitée pour des v.a. à valeurs dans $\mathbb{R}^d$. Pour un traitement très élégant et rapide dans le cas de v.a. réelles, voir le chapitre 16 de Williams, *Probability with martingales*. On a besoin de v.a. vectorielles à cause, notamment, du traitement des vecteurs gaussiens.

2.1 Généralités

Définition 2.1. *Soit μ une mesure finie sur $\mathbb{R}^d$. On appelle transformée de Fourier de μ l'application de $\mathbb{R}^d$ dans $\mathbb{C}$ définie par*

$$(\mathcal{F}\mu)(t) = \hat{\mu}(t) = \int_{\mathbb{R}^d} e^{ix.t} \mu(dx).$$

(Ici, $x.t$ est le produit scalaire usuel de x et $t \in \mathbb{R}^d$)
On appelle fonction caractéristique d'une variable aléatoire X la transformée de Fourier de sa loi $\mu = \mathbb{P}_X$. Elle est notée φ_X et vaut

$$\varphi_X(t) = \mathbb{E}(e^{iX.t}) = \hat{\mu}(t), \quad (t \in \mathbb{R}^d).$$

.

Proposition 2.1. ***Propriétés élémentaires***
1) $\hat{\mu}(0) = \mu(\mathbb{R}^d)$, $\varphi_X(0) = 1$.
2) $\forall t \in \mathbb{R}^d$, $|\hat{\mu}(t)| \leq \mu(\mathbb{R}^d)$, $|\varphi_X(t)| \leq 1$.
3) $\forall t \in \mathbb{R}^d$, $\hat{\mu}(-t) = \overline{\hat{\mu}}(t)$, $\varphi_X(-t) = \overline{\varphi_X(t)}$.
4) Soient $A \in \mathcal{L}(\mathbb{R}^d, \mathbb{R}^k)$ et $b \in \mathbb{R}^k$, alors

$$\forall t \in \mathbb{R}^k, \qquad \varphi_{AX+b}(t) = \varphi_X(A^\star t) e^{ib.t},$$

où $A^\star$ désigne l'adjoint de A.

On va montrer pour deux mesures bornées μ_1, μ_2 que
- $\hat{\mu}_1 = \hat{\mu}_2$ implique $\mu_1 = \mu_2$.

- $[\mathcal{F}(\mu_1 \otimes \mu_2)](t_1, t_2) = [\mathcal{F}\mu_1](t_1)[\mathcal{F}\mu_2](t_2)$.

Deux v.a. qui ont même transformée de Fourier sont égales (injectivité).
Pour deux v.a. aléatoires indépendantes X_1 et X_2 :

$$\varphi_{(X_1,X_2)}(t_1,t_2) = \varphi_{X_1}(t_1)\varphi_{X_2}(t_2), \quad \varphi_{X_1+X_2}(t) = \varphi_{X_1}(t)\varphi_{X_2}(t)$$

Exercice 2.1. Montrer que les fonctions $\hat{\mu}$ et φ_X sont uniformément continues sur $\mathbb{R}^d$.
Indication : Fixer $\varepsilon > 0$ et choisir n tel que $\mu(B(0,n)^c) \leq \varepsilon$, montrer et utiliser l'inégalité $|e^{ix.u} - e^{ix.t}| \leq ||u-t||\ ||x||$

Exercice 2.2. Exemple très important : la gaussienne
On pose

$$g_\sigma(x) = \frac{1}{(\sigma\sqrt{2\pi})^d} e^{-\frac{||x||^2}{2\sigma^2}}.$$

Vous devez savoir démontrer sans aucune hésitation les relations suivantes :

1) La fonction g_d est une densité de probabilité sur $\mathbb{R}^d$.

2) Pour tout $f \in C_b(\mathbb{R}^d)$, $\lim_{\sigma\to 0}(f * g_\sigma)(x) = f(x)$.

3) Montrer (de trois facons) que pour tout $t \in \mathbb{R}^d$,

$$\hat{g}_1(t) = e^{-\frac{||t||^2}{2}} = (\sqrt{2\pi})^d g_1(t).$$

Solutions :

1) Se ramener au cas de la dimension 1 avec Fubini et et se souvenir comment l'on montre que $\int_{\mathbb{R}} e^{-x^2/2}dx = \sqrt{2\pi}$. Il faut pour cela regarder $\int_{\mathbb{R}^2} \exp(-[x^2+y^2]/2)dxdy$ en polaire

2) Utiliser le théorème de Lebesgue après changement de variable $y = \sigma z$.

3) On commence à se ramener au calcul en dimension 1 par le théorème de Fubini.
- *Première méthode très simple que l'on peut trouver dans le chapitre 9 de Bony,* Cours d'Analyse, *sur la transformée de Fourier. Poser $g(s) = \int_{\mathbb{R}} e^{isx}e^{-x^2}dx$ et montrer en utilisant le théorème de dérivation sous le signe somme que g est C^1 et déduire de l'expression de g' en intégrant par parties que g est solution de l'équation différentielle $2g'(s) + sg(s) = 0$.*
- *Une autre méthode utilise la formule de Cauchy appliquée à la fonction holomorphe $g(z) = \exp(-z^2/2)$ sur le circuit composé des segments $[(-r,0),\ (r,0)]$, $[(r,0),\ (r,t)]$, $[(r,t),\ (-r,t)]$, $[(-r,t),\ (-r,0)]$ en faisant tendre r vers l'infini.*
- *Une troisième méthode élégante consiste à remarquer que*

$$\hat{g}_1(t) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-\frac{(x-it)^2}{2}} e^{-\frac{t^2}{2}} dx.$$

On pose alors

$$f(z) := \int_{\mathbb{R}} e^{-\frac{(x-z)^2}{2}} dx.$$

On démontre très soigneusement en utilisant le théorème de dérivation sous le signe somme que f est holomorphe sur $\mathbb{C}$. Comme f est constante et égale à $\sqrt{2\pi}$ sur $\mathbb{R}$ par un changement de variable trivial, on conclut par prolongement analytique.

2.2 Injectivité de la transformée de Fourier

Définition 2.2. *(Convolution d'une mesure bornée et d'une fonction). Soit μ une mesure bornée sur $\mathbb{R}^d$ et f une fonction borélienne telle que, pour tout x, la fonction $y \to f(y-x)$ soit μ-intégrable. On pose*

$$(f * \mu)(x) = \int_{\mathbb{R}^d} f(x-y)\mu(dy).$$

Exercice 2.3. Théorème d'injectivité de la transformée de Fourier sur les mesures bornées
Il s'agit de montrer que *deux mesures bornées sur $\mathbb{R}^d$ qui ont même transformée de Fourier sont égales.* Cette démonstration suit Ouvrard, Tome II, pp. 202-204.

1) Soit μ une mesure bornée sur $\mathbb{R}^d$. Montrer en appliquant le théorème de Fubini et la relation (d) de l'exercice précédent que

$$(g_\sigma * \mu)(y) = \frac{1}{(\sqrt{2\pi})^d} \int_{\mathbb{R}^d} \hat{\mu}(v) g_1(\sigma v) e^{-iy.v} dv.$$

2) Expliquer pourquoi une mesure bornée sur $\mathbb{R}^d$ est déterminée par la donnée pour tout $f \in C_K(\mathbb{R}^d)$ des intégrales $\int_{\mathbb{R}^d} f d\mu$.
En déduire qu'il suffit de donner une expression de $\int f d\mu$ en fonction de $\hat{\mu}$ pour montrer l'unicité.

3) Montrer en appliquant le théorème de Lebesgue que

$$\int_{\mathbb{R}^d} f d\mu = \lim_{\sigma \to 0} \int_{\mathbb{R}^d} (f * g_\sigma)(x) d\mu(x).$$

4) En appliquant le théorème de Fubini, en déduire que

$$\int_{\mathbb{R}^d} f d\mu = \lim_{\sigma \to 0} \int_{\mathbb{R}^d} f(y)(g_\sigma * \mu)(y) dy.$$

5) En utilisant le 1), déduire que

$$\int_{\mathbb{R}^d} f d\mu = \frac{1}{(\sqrt{2\pi})^d} \lim_{\sigma \to 0} \int f(y) \left(\int \hat{\mu}(v) g_1(\sigma v) e^{-iy.v} dv \right) dy.$$

6) Enoncer le théorème d'injectivité de la transformée de Fourier.

7) Application : soit μ une mesure bornée sur $\mathbb{R}^d$ telle que $\hat{\mu}$ soit Lebesgue-intégrable. Alors μ est absolument continue par rapport à la mesure de Lebesgue et sa densité est donnée par

$$h(x) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} \hat{\mu}(t) e^{-ix.t} dt.$$

Indication :
1) On utilise $g_1\hat{(\sigma.)}(y-x) = g_1\hat{(\sigma.)}(x-y) = \sqrt{2\pi} g_\sigma(x-y)$ *pour*

$$\int_{\mathbb{R}^d} \hat{\mu}(v) g_1(\sigma v) e^{-iy.v} dv = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} e^{ix.v} \mu(dx) e^{-iy.v} dv = \int_{\mathbb{R}^d} \mu(dx) g_1\hat{(\sigma.)}(y-x)$$

2) Il suffit de montrer que ces mesures coïncident sur les ouverts car ils forment un π-système des boréliens. On montrera brièvement que toute fonction caractéristique d'ouvert est une limite croissante de fonctions continues à support compact (En effet, soit $g_r(x) = \max(0, 1 - r||x||)$. Quand $r \to 0$, cette fonction tend en croissant vers 1. Posant $h_r(x) = \inf(1, \frac{1}{r} d(x, \Omega^c))$, puis $k_r(x) = g_r(x) h_r(x)$, on vérifie que cette suite de fonctions convient, quand $r \to 0$).
7) Commencer par montrer en appliquant le théorème de Lebesgue que

$$\int_{\mathbb{R}^d} \hat{\mu}(v) g_1(\sigma v) e^{-iy.v} dv \to \frac{1}{(\sqrt{2\pi})^d} \int_{\mathbb{R}^d} \hat{\mu} e^{-iy.v} dv$$

Puis, appliquer à nouveau le théorème de Lebesgue dans la relation 5) avec $f \in C_K(\mathbb{R}^d)$.

Exercice 2.4. Transformée de Fourier et indépendance (Ouvrard, tome II, pp. 204-208).
On considère des mesures bornées μ_1 et μ_2 sur $\mathbb{R}^{d_1}$ et $\mathbb{R}^{d_2}$ et la mesure produit $\mu_1 \otimes \mu_2$.

1) Montrer en appliquant le théorème de Fubini (on rappelle que $\mathcal{F}$ désigne la tranformation de Fourier) que

$$[\mathcal{F}(\mu_1 \otimes \mu_2)](t_1, t_2) = [\mathcal{F}\mu_1](t_1)[\mathcal{F}\mu_2](t_2).$$

2) Soit $X = (X_1, X_2)$ une v.a. à valeurs dans $\mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$. Déduire du 1) que X_1 et X_2 sont indépendantes si et seulement si

$$\varphi_{(X_1,X_2)}(t_1, t_2) = \varphi_{X_1}(t_1)\varphi_{X_2}(t_2).$$

On utilisera l'injectivité de la transformée de Fourier.

3) Remarquer que $\varphi_{X_1}(t_1) = \varphi_{X_1,X_2}(t_1, 0)$.

4) Si μ_1 et μ_2 sont des mesures bornées sur $\mathbb{R}^d$, on a $\mathcal{F}(\mu_1 * \mu_2) = \mathcal{F}\mu_1 \mathcal{F}\mu_2$.
Rappel (Ouvrard p. 63) : la convolée de deux mesures μ_1 et μ_2, $\mu_1 * \mu_2$ est la mesure image de $\mu_1 \otimes \mu_2$ par l'application somme. Cela signifie que pour toute fonction mesurable positive,

$$\int_{\mathbb{R}^d} f \mu_1 * \mu_2 = \int_{\mathbb{R}^d \times \mathbb{R}^d} f(x_1 + x_2) \mu_1 \otimes \mu_2(dx_1 dx_2).$$

Appliquer cette formule avec $f(x) = e^{ix.t}$ et utiliser le théorème de Fubini.

Exercice 2.5. Un contrexemple (Brémaud P., *Introduction aux probabilités : Modélisation des phénomènes aléatoires*, p. 134, exercice 13).
On considère le sous-ensemble de $\mathbb{R}^2$,

$$C = \{(x, y) \in [0, 1]^2,\ x - \frac{1}{2} \leq y \leq x \text{ ou } y \geq x + \frac{1}{2}.\}$$

1) Dessiner C.

2) Soit X et Y deux variables aléatoires de densité de probabilité $f_{X,Y}$ constante sur C, nulle ailleurs. Calculer la valeur de la constante et les densités de X et Y. En déduire que X et Y ne sont pas indépendantes.

3) Calculer les fonctions caractéristiques de X, Y et $X+Y$ pour montrer que

$$\varphi_X(t)\varphi_Y(t) = \varphi_{X+Y}(t).$$

Solution

Un calcul élémentaire montre que la constante vaut 2. On désigne par f_X la densité de X et par f_Y celle de Y. On a

$$f_X(x) = \int_{-\infty}^{+\infty} f_{X,Y}(x,y)dy = \mathbb{1}_{[0,1]}(x)$$

et

$$f_Y(y) = \mathbb{1}_{[0,1]}(y).$$

X et Y ne sont pas indépendantes puisque

$$f_X(x)f_Y(y) \neq f_{X,Y}(x,y).$$

Or on a

$$\varphi_{X+Y}(u) = \left(\frac{e^{iu}-1}{iu}\right)^2, \ \varphi_X(u) = \varphi_Y(u) = \left(\frac{e^{iu}-1}{iu}\right).$$

Exercice 2.6. (Williams, *Probability with martingales* p. 238) Calculer la fonction caractéristique de la loi uniforme sur $[-1,1]$. Montrer qu'il n'existe pas de variables aléatoires i.i.d. X et Y telles que $X-Y$ ait cette loi uniforme sur $[-1,1]$.

2.3 Fonctions caractéristiques : ce qu'il faut absolument savoir

1) La défénition d'une fonction caractérisque.

2) Retrouver rapidement les propriétés élémentaires sur les fonctions caractéristiques.

3) Connaître par coeur la fonction caractéristique de la Gaussienne et connaître une méthode pour la retrouver. Savoir aussi comment calculer les fonctions caractéristiques des lois usuelles.

4) Savoir que la fonction caractéristique caractérise complètement la loi d'une v.a. (injectivité de la transformée de Fourier) .

5) Etre au courant de la formule d'inversion de Levy. Aller y jeter un oeil, ou même les deux ([**?**], pp. 175-177).

6) Savoir que si X et Y sont indépendantes alors la fonction caractéristique de (X,Y) s'écrit comme le produit direct des fonctions caractéristiques de X et Y et réciproquement.

7) Connaître le théorème de Lévy.

8) Savoir que si X possède un moment d'ordre n alors la fonction caractéristique de X est n-fois dérivable et le théorème de dérivation sous le signe somme s'applique. Etre au courant aussi que la réciproque est fausse : On peut avoir une fonction caractéristique dérivable en 0 sans que la variable aléatoire X ait un moment d'ordre 1 (Voir ouvrard tome 2 par exemple).

3 Convolution de mesures

La convolution de deux objets (fonctions/mesures, en discter ou en continu) en x est donné par $\int f(x-y)g(y)$

Définition 3.1. *(Convolution de deux fonctions). Soit f_1 et f_2 deux fonctions boréliennes intégrables sur de $\mathbb{R}^d$ dans $\mathbb{R}$. On pose*

$$(f_1 * f_2)(x) = \int_{\mathbb{R}^d} f_1(x-y)f_2(y)dy.$$

Définition 3.2. *(Convolution d'une mesure bornée et d'une fonction). Soit μ une mesure bornée sur $\mathbb{R}^d$ et f une fonction borélienne telle que, pour tout x, la fonction $y \to f(y-x)$ soit μ-intégrable. On pose*

$$(f * \mu)(x) = \int_{\mathbb{R}^d} f(x-y)\mu(dy).$$

La convolée de deux mesures μ_1 et μ_2, $\mu_1 * \mu_2$ est la mesure image de $\mu_1 \otimes \mu_2$ par l'application somme. Cela signifie que pour toute fonction mesurable positive,

$$\int_{\mathbb{R}^d} f\mu_1 * \mu_2 = \int_{\mathbb{R}^d \times \mathbb{R}^d} f(x_1+x_2)\mu_1 \otimes \mu_2(dx_1 dx_2).$$

Plus spécifiquement, on peut définir la convolée de deux mesures sur $\mathbb{R}$, μ_1 et μ_2, par

$$\mu_1 * \mu_2(]-\infty, x]) = \int_{-\infty}^{\infty} \mu_1(]-\infty, x-y])\mu_2(dy).$$

Proposition 3.1. *Si X et Y sont deux variables aléatoires indépedantes de lois respectives μ et λ, alors la loi de $X+Y$ est la convolée de la loi de λ et μ.*

Ce qui fait que la loi de la somme de n variables i.i.d. est la convolée $n$ième de la loi commune. Nous explorons dans le problème suivant la question de la répartition des points qui se trouvent à des distances i.i.d. Les convolées successives jouent un rôle primordial dans l'approche fonctionnelle de ce problème, dont les applications en modélisation et en statistiques sont multiples. Nous donnons par ici une conséquance naturelle (le théorème de renouvellement) et une conséquence paradoxale (le paradxe de l'autobus)

3.1 Problème : Théorie du renouvellement

Ce problème mélange différentes notions d'analyse : convolution, point fixe, uniforme continuité, densité, espace de fonctions... Il est motivé par des questions de probabilités abordées à la fin. Il s'inspire beaucoup de Feller tome II *Introduction to probability.* Parmi les applications possibles, on donne à la fin la loi du temps d'attente du prochain bus passant à une station où les temps d'interarrivées des bus sont i.i.d. : elle fait apparaitre un paradoxe flagrant puisque le temps moyen d'attente n'est pas le temps moyen entre deux bus divisé par 2, mais peut même être supérieur au temps moyen entre deux bus.

On considère une mesure positive λ sur les boréliens de $\mathbb{R}$ et on note supp λ son support :

$$\text{supp } \lambda := \{x \in \mathbb{R} : \forall \epsilon > 0,\ \lambda([x-\epsilon, x+\epsilon]) > 0\}$$

La convolution entre deux mesures positives λ et μ est notée $\lambda \star \mu$ et définie (quand elle existe) par

$$\lambda \star \mu(]-\infty, x]) = \int_{-\infty}^{\infty} \lambda(]-\infty, x-y])\mu(dy).$$

On va s'intéresser dans ce problème aux convolées successives d'une *mesure de probabilité λ à support dans* $\mathbb{R}^+$:

$$\lambda(]-\infty, 0[) = 0, \quad \lambda([0, \infty)) = 1$$

et l'on note $\lambda^{\star k}$ la suite de mesure définie par

$$\lambda^{\star 0} = \delta_0, \quad \lambda^{\star k+1} = \lambda \star \lambda^{\star k}$$

où δ_0 est la masse de Dirac en 0. Plus précisement, on considère la mesure positive

$$U = \sum_{k=0}^{\infty} \lambda^{\star k}$$

qui est la clef pour comprendre la répartition asymptotique des points disposés à intervalles indépendants identiquement distribués.

3.2 Densité à l'infini

1) Vérifier que supp $U = \cup_{k=0}^{\infty}$supp $\lambda^{\star k}$.

2) Montrer que supp U contient 0 et est stable par addition, i.e. $0 \in$ supp U et $\forall a, b \in$ supp $U, a+b \in$ supp$U$.

Le support de U est un ensemble qui tend à se densifier quand x augmente et on va déterminer dans quel cas il devient dense à l'infini. On dit que $D \subset [0, \infty[$ est dense à l'infini ssi

$$\forall \epsilon > 0, \quad \exists x_0 \geq 0, \quad \forall x \geq x_0 : \quad D \cap [x, x+\epsilon] \neq \varnothing.$$

On pose $r = \inf\{b - a \ : \ a < b,\ a, b \in \text{supp } U\}$.

3) Montrer que si $r > 0$, alors supp $U \subset r\mathbb{N}$.

4) Montrer que si $r = 0$ alors supp U est dense à l'infini.

5) Montrer que si $\{1, \sqrt{2}\} \subset$ supp λ, alors supp$U$ est dense à l'infini.

En conclusion, ou bien supp U est dense à l'infini, ou bien il est inclus dans un réseau. On se place désormais dans le premier cas, c'est à dire que pour tout $r > 0$, supp $U \subsetneq r\mathbb{N}$, et on dit que λ n'est pas *arithmétique*.

3.3 Equation de convolution

On définit la convolution $f \star \mu$ entre une fonction $f : \mathbb{R} \to \mathbb{R}$ et une mesure positive μ comme (quand elle existe)

$$f \star \mu(x) = \int_{-\infty}^{\infty} f(x-y)\mu(dy).$$

On cherche ici à résoudre l'equation

$$f = g + f \star \lambda \qquad (R)$$

où g est une fonction mesurable par rapport aux boréliens de $\mathbb{R}$, bornée et à support compact inclus dans $\mathbb{R}^+$ et λ est une mesure de probabilité à support dans $\mathbb{R}^+$ qui n'est pas arithmétique (voir la question précédente). Pour cela on introduit la mesure U_n définie par

$$U_n = \sum_{k=0}^{n} \lambda^{\star k}.$$

1) Montrer que pour tout $h \geq 0$, il existe $C \geq 0$ tel que pour tout $x \geq 0$, $U([x, x+h]) \leq C$.

Ce résultat dit simplement que l'intensité du nuage de points n'explose pas quand $x \to \infty$.

2) Montrer que $f_n = g \star U_n$ converge simplement vers $g \star U$ et est uniforment bornée en n sur $\mathbb{R}$.

3) Montrer que si $\zeta : \mathbb{R} \to \mathbb{R}$ est bornée, uniformement continue et vérifie $\zeta = \zeta \star \lambda$ et $\forall x \in \mathbb{R}$: $\zeta(x) \leq \zeta(0)$, alors $\forall x \in \mathbb{R} : \zeta(x) = \zeta(0)$.

4) Montrer que l'équation (R) admet une unique solution bornée sur tout compact qui est égale à $g \star U$.

5) On suppose ici g continue à support dans $[0, K]$ et on note $f = g \star U$ la solution considérée dans la précédente question.

a) Montrer que f est bornée et uniformement continue.

b) En supposant que g est $\mathcal{C}^1$ à support dans $[0, K]$, montrer que f est dérivable puis que $\limsup_{x\to\infty} f'(x) = 0$ en considérant $\zeta_n(.) = f'(t_n + .)$ avec $\lim_{n\to\infty} f'(t_n) = \limsup_{x\to\infty} f'(x)$.

En déduire que pour tout $a \geq 0$, $f(x+a) - f(x) \to 0$ quand $x \to \infty$.

c) Montrer que pour tout $a \geq 0$, $f(x+a) - f(x) \to 0$ quand $x \to \infty$.

3.4 Théoremes de renouvellement

1) On note $\mathcal{C}_K(\mathbb{R})$ l'ensemble des fonctions continues sur $\mathbb{R}$ à support compact que l'on munit de la norme infinie sur $\mathbb{R}$. On dit qu'une suite de mesures ν_n converge faiblement vers ν quand pour tout $f \in \mathcal{C}_K(\mathbb{R})$,

$$\int_{-\infty}^{\infty} f(x)\nu_n(dx) \overset{n\to\infty}{\longrightarrow} \int_{-\infty}^{\infty} f(x)\nu(dx).$$

Soit ν_n des mesures positives sur $\mathbb{R}$ telles que pour tout pour tout compact K, $\sup_{n\in\mathbb{N}} \nu_n(K) < \infty$.

a) Montrer qu'il existe une sous suite extraite $\nu_{\phi(n)}$ qui converge faiblement vers une mesure positive ν.

b) Montrer que si ν_n converge faiblement vers la mesure de Lebesgue dx, que f est intégrable par rapport à ν_n et directement Riemann intégrable sur $\mathbb{R}$ (voir la définition plus bas) alors,

$$\int_{-\infty}^{\infty} f(x)\nu_n(dx) \overset{n\to\infty}{\longrightarrow} \int_a^b f(x)dx.$$

2) Montrer que si une mesure positive ν est finie sur tout compact vérifie que pour tous $z \in \mathbb{R}$ et $f \in \mathcal{C}_K(\mathbb{R})$, $\int_{-\infty}^{\infty} f(x)\nu(z+dx) = \int_{-\infty}^{\infty} f(x)\nu(dx)$, alors ν est proportionelle à la mesure de Lebsegue.

3) On peut maintenant démontrer le résultat fondamental de renouvellement, pour toute mesure λ non arithmétique. Ce résultat indique que quand $x \to \infty$, le nombre de points moyen du nuage qui se trouve dans $[x, x+h]$ devient proportionnel à h.
a) Montrer que si λ n'est pas arithmétique, pour toute fonction f Riemann directement intégrable sur $\mathbb{R}^+$,

$$\int_0^{\infty} f(x-y)U(dy) \overset{x\to\infty}{\longrightarrow} m^{-1}\int_0^{\infty} f(y)dy$$

En particulier, pour tout $h > 0$, $U([x, x+h]) \overset{x\to\infty}{\longrightarrow} h/m$.
b) Montrer que ce dernier résultat est faux si λ est arithmétique.

3.5 Interprétation probabiliste

On considère maintenant un espace de probabilité $(\Omega, \mathbb{P})$.
1) Pour tous λ, μ mesures de probabilité, vérifier que $\lambda \star \mu = \mu \star \lambda$ puis que $\lambda \star \mu$ est la loi de la somme de deux variables aléatoires indépendantes de loi respectives λ et μ.

Dorénavant, λ est une mesure dont le support est inclus dans $[0, \infty[$ et de moyenne finie : $\int_0^{\infty} x\lambda(dx) < \infty$.
On considère $(X_i : i \geq 1)$ des v.a. sur $(\Omega, \mathbb{P})$ indépendantes identiquement distribuées de loi commune λ. On s'intéresse à l'ensemble aléatoire discret

$$\mathcal{E}(\omega) := \{\sum_{i=1}^{k} X_i(\omega) : k \geq 0\}.$$

C'est à dire que à chaque ω on associe un ensemble $\mathcal{E}(\omega)$ tel que la distance entre le $i$ème et le $i+1$ème point est égale à X_i.

On note

$$m = \int_0^{\infty} x\lambda(dx) = \mathbb{E}(X_1), \quad N_x = \sup\{k : \sum_{i=1}^{k} X_i \leq x\}$$

On commence par un résultat simple qui dit que le nombre de points du nuage dans l'intervalle $[0, x]$ croit linéairement en x avec une vitesse inverse à m.
2) Montrer que $N_x/x \to 1/m$ p.s. quand $x \to \infty$.

L'étude plus fine de ce nuage de points repose sur son intensité donnée par la mesure U définie au début. En effet λ^k donne la loi de l'abscisse du $k$ème points de $\mathcal{E}$, $U([a, b])$ donne le nombre moyen de points de $\mathcal{E}$ entre a entre b.

3) Montrer que pour tout $0 \leq a \leq b$, $U([a,b]) = \mathbb{E}(\#\{z(\omega) \in \mathcal{E} \cap [a,b]\})$.

Le résultat 3.3.a) indique donc que si la loi de X n'est pas arithmétique, alors le nombre moyen de points de $\mathcal{E}$ entre x et $x+t$ converge vers t/m. Ce qui est très naturel mais tout de même faut si la loi de X est arithmétique et aboutit au paradoxe suivant.

3.6 Application au paradoxe de l'autobus

On suppose toujours que la loi de X n'est pas arithmétique et on considère maintenant que les points de $\mathcal{E}$ sont les instants de passage d'un bus à une station. Ceci signifie que les temps successifs entre deux bus sont indépendants et identiquement distribués comme la variable aléatoire X.

Lorsqu'un passager arrive à l'instant $t \geq 0$, le bus suivant arrive lui à l'instant

$$T_t = \inf\{x \geq t : x \in \mathcal{E}\}$$

et il attend donc un temps

$$A_t = T_t - t.$$

1) Montrer que pour tout $x \geq 0$,

$$\mathbb{P}(A_t \leq x) \overset{t\to\infty}{\longrightarrow} m^{-1} \int_0^x \mathbb{P}(X \geq y)dy.$$

2) Montrer que si $\mathbb{E}(X^2) < \infty$, alors

$$\mathbb{E}(A_t) \overset{t\to\infty}{\longrightarrow} \frac{\mathbb{E}(X^2)}{2\mathbb{E}(X)}$$

Définition d'une fonction directement Riemann intégrable sur I C'est une notion plus forte que Riemann intégrable, même si elles coincident sur un segment. Pour un recouvrement I par des intervalles disjoints

$$I = \sqcup_{i\in\mathbb{N}} I_n$$

on définit les petites et grandes sommes de Darboux par

$$d(f,(I_i : i \in \mathbb{N})) = \sum_{i\in\mathbb{N}} |I_i| \inf_{x\in I_i} f(x), \qquad D(f,(I_i : i \in \mathbb{N})) = \sum_{i\in\mathbb{N}} |I_i| \sup_{x\in I_i} f(x).$$

On dit alors que f est directement Riemann intégrable sur I d'intégrale $\int_I f(x)dx$ quand pour toute suite de recouvrement $((I_i^n : i \in \mathbb{N})n \in \mathbb{N})$ dont le pas tend vers zéro :

$$\sup_{i\in\mathbb{N}} |I_i^n| \overset{n\to\infty}{\longrightarrow} 0$$

les sommes de Darboux convergent vers la même valeur, i.e.

$$D(f,(I_i^n : i \in \mathbb{N})) - d(f,(I_i^n : i \in \mathbb{N})) \overset{n\to\infty}{\longrightarrow} 0$$

et on note alors

$$\int_I f(x)dx = \lim_{n\to\infty} D(f,(I_i^n : i \in \mathbb{N})) = \lim_{n\to\infty} d(f,(I_i^n : i \in \mathbb{N})).$$

Par les méthodes usuelles, on vérifie que cette définition est bien indépendante de la suite de recouvrement.

INDICATIONS.
1.2) Montrer d'abord que supp $\mu_1 +$ supp $\mu_2 \subset$ supp $\mu_1 \star \mu_2$.
1.3) Poser $h_0 = b - a \in$ supp U où $a < b$ et $r \leq h_0 < 2r$. Faire un dessin en remarquant que les points de la forme $na + k(b-a) \in$ supp U pour $0 \leq k \leq n$. Montrer que $h_0 = r$ et supp $U \subset h_0\mathbb{N}$.
1.4) Faire un dessin. Soit $\epsilon > 0$, il existe $a, b \in$ supp U tel que $a < b \leq a + \epsilon$. Considérer alors $n \in \mathbb{N}$ tel que $(n+1)a \leq nb$.
2.1) Considérer $U_n([0,x]) - \lambda \star U_n([0,x])$.
2.3) En utilisant $\zeta = \zeta \star \lambda^k$, montrer que $\zeta(-x) = \zeta(0)$ pour $x \in$ supp U puis $\zeta(-x) \to \zeta(0)$ quand $x \to \infty$.
3.1.a) Utiliser qu'il existe une suite de fonctions dense dans $\mathcal{C}_K(\mathbb{R})$ (séparabilité). Faire alors une extraction diagonale et utliser le théorème de Riesz.
3.1.b) Encadrer une fonction Riemann intégrable sur un segment par ses sommes de Darboux.
3.2) Construire une suite de fonctions qui converge en croissant vers la fonction indicatrice d'un segment.
4.1) Utiliser Fubini avec $\lambda(]-\infty, x]) = \int_{-\infty}^{x} \lambda(y)$.
4.2) Utiliser la loi des grands nombres.

CORRECTION.

1.3) On pose $h_0 = b - a$ avec $a < b \in$ supp U et et $r \leq h_0 < 2r$. . Introduire ensuite $n \geq 0$ tel que $(n+1)a \leq nb$. Alors $(n+1)a \in [na, nb]$ et il existe $0 \leq k \leq n$ tel que $|(n+1)a - (na + kh_0)| < r$. Or $(n+1)a, (na + kh_0) \in$ supp U par additivité de supp U. Donc $(n+1)a = na + kh_0$. Ceci implique $a = kh_0$ et $b = (k+1)h_0$. On en déduit de même que pour tout $z \in$ supp U, il existe $n \in \mathbb{N}$ tel que $|z - nh_0| = |z + a - (a + nh_0)| < r$ et donc $z = nh_0$ puisque $z + a, (a + nh_0) \in$ supp U. D'où supp $U \subset h_0\mathbb{N}$.

2.1) Pour tout $n \in \mathbb{N}$ et $x \geq 0$,

$$U_n([0,x]) - \lambda \star U_n([0,x]) = \int_0^x \lambda([x-y, \infty[) U_n(dy) = 1 - \lambda^{\star n+1}([x, \infty[)$$

En posant $\tau > 0$ tel que $\lambda([\tau, \infty[) > 0$, cette identité entraine que pour tous $n \in \mathbb{N}$ et $x \geq 0$, $\lambda([\tau, \infty[) U_n([x - \tau, x]) \leq 1$. En faisant tendre $n \to \infty$, on a pour tout $x \geq 0$:

$$U([x - \tau, x]) \leq 1/\lambda([\tau, \infty[).$$

Le résultat est alors en obtenu en écrivant $[x, x+h] \subset \cup_{i=0}^{n(h)} [x + i\tau, x + (i+1)\tau]$ avec $n(h) = \inf\{n \in \mathbb{N} : (n+1)\tau \geq h\}$.

2.3) Pour tout $k \in \mathbb{N}$, $\zeta = \zeta \star \lambda^k$ et donc $\int_0^\infty [\zeta(0) - \zeta(-x)] \lambda^{\star k}(dx) = 0$. En utilisant la continuité de ζ et le fait que pour tout $x \geq 0$, $\zeta(-x) \leq \zeta(0)$, on obtient que $\zeta(-x) = \zeta(0)$ pour tout $x \in \text{supp}\lambda^{\star k}$ et donc pour tout $x \in$ supp U.
Comme supp U est dense à l'infini et ζ est uniformement continue, $\zeta(-x) \to \zeta(0)$ quand $x \to \infty$.

Fixons alors $x \geq 0$ et notons M un majorant de $|\zeta|$. Pour tout $A \geq 0$,

$$\begin{aligned} |\zeta(x) - \zeta(0)| &= \int_0^\infty |\zeta(x-y) - \zeta(-y)| \lambda^{\star k}(dy) \\ &\leq 2M\lambda^k([0,A]) + \sup\{|\zeta(x-y) - \zeta(-y)| : y \geq A\}. \end{aligned}$$

Pour tout $\epsilon > 0$, on peut choisir $A \geq 0$ tel que $\sup\{|\zeta(x-y) - \zeta(-y)| : y \geq A\} \leq \epsilon$. Grace à 2.1), on peut ensuite trouver k tel que $\lambda^k([0,A]) \leq \epsilon$. Ceci implique que $\zeta(x) = \zeta(0)$.

2.4) On pose $f = g \star U$. Grâce à 2.2), $f_n \to f$ simplement et par convergence dominée, $f_n \star \lambda \to f \star \lambda$ simplement. En utilisant $f_{n+1} = g + f_n \star \lambda$, on obtient que $f \star \lambda$ est solution de (R).
Pour l'unicité, on remarque que si $f = f \star \lambda$, alors pour tout $k \geq 1$, $f = f \star \lambda^{\star k}$. On conclut grace à $|f(x)| \leq \sup_{[0,x]} |f|.\lambda^{\star k}([0,x])$ puisque le deuxième terme tend vers 0 quand $k \to \infty$ par 2.1).

2.5.a) En remarquant que $|f(x)| = |g \star U(x)| \leq U([x-K,x]) \sup |g|$, on obtient que f est bornée sur $\mathbb{R}$ grace à 2.1). Comme pour tout $a \geq 0$, $f(a+.) - f(.) = [g(a+.) - g(.)] \star U$, pour tout $x \in \mathbb{R}$,

$$|g(x+a) - g(x)| \leq U([x-K, x+a]) \sup_{y \in \mathbb{R}} |g(x+a-y) - g(x-y)|$$

et le membre de droite tend vers 0 uniformement en x quand $a \to 0$ en utilisant 2.1) et le théorème de Heine qui assure l'uniforme continuité de g.

2.5.b) Grace à 2.1), on peut appliquer le théorème de dérivation sous le signe inétgrale et $f' = g' \star U = g' + f' \star \lambda$. Comme g' est continue à support compact, la question précédente assure que f' est bornée et uniformement continue.
On note $L = \limsup_{x \to \infty} f'(x)$ et on introduit une suite $t_n \to \infty$ tel que $\lim_{n \to \infty} f'(t_n) = L$. Considérons la suite de fonctions sur $\mathbb{R}$: $\zeta_n(.) = f'(t_n + .)$. $(\zeta_n : n \in \mathbb{N})$ est une famille bornée de fonctions en norme infinie car f' bornée. De plus c'est une famille equicontinue par uniforme continuité de f'. Donc, par le théorème d'Ascoli Arzela, on peut extraire une sous suite qui converge uniforment sur un compact donné. Par extraction diagonale, on peut extraire une suite qui converge uniformement vers ζ sur tout compact de $\mathbb{R}$.
Or $\zeta_n = g(t_n + .) + \zeta_n \star \lambda$ implique que $\zeta = \zeta \star \lambda$. De plus, pour tout $x \in \mathbb{R}$, $\zeta(x) \leq L = \zeta(0)$ et on peut appliquer 2.4). On obtient que pour tout $x \in \mathbb{R}$, $\zeta(x) = L$ et donc $f'(t_n + x) \to L$ quand $n \to \infty$.
Alors pour tout $a \geq 0$, $f(t_n + a) - f(t_n) = \int_0^a f'(t_n + x)dx \to aL$ quand $n \to \infty$ par convergence dominée ou convergence uniforme sur un compact. Mais f bornée donc $L = 0$.

Par symétrie, on obtient que $\lim_{x \to \infty} f'(x) = 0$ et de même pour tout $a \geq 0$, $f(x+a) - f(x) = \int_0^a f'(x+y)dy \to 0$ par convergence dominée.

2.5.c) On considère une suite g_n de fonctions $\mathcal{C}^1$ à support dans $[0, K+1]$ qui converge uniformement vers g. Elle peut etre obtenue par exemple facilement à partir de g grace à un noyau régularisant.
Alors $f_n = g_n \star U$ converge uniformément vers $f = g \star U$ en utilisant 2.1) et

$$|g_n \star U(x) - g \star U(x)| \leq U([x-K-1, x]) \parallel g_n - g \parallel_\infty .$$

Le résultat se dérive alors directement de la question précédente.

3.1.a) On construit d'abord une suite de fonctions continues denses dans les fonctions continues à support dans $[-N,N]$ où $N \in \mathbb{N}$. Pour cela, on peut utiliser la densité de la suite formée par les polynomes à coefficients rationnels dans les fonctions continues pour la norme uniforme sur $[-N,N]$. On obtient alors une suite dense $(f_k : k \in \mathbb{N})$ dans $\mathcal{C}_K(\mathbb{R})$ en faisant une extraction diagonale à partir de ces suites pour $N \in \mathbb{N}$.
Or par hypothèse, pour tout segment I, $\sup_{n\in\mathbb{N}} \nu_n(I) < \infty$. Donc pour tout $k \in \mathbb{N}$, la suite $(\int_{-\infty}^{\infty} f_k(x)\nu_n(dx) : n \in \mathbb{N})$ est bornée. Elle admet donc une sous suite convergente et par extraction diagonale, il existe une sous suite $\widetilde{\nu}_n$ telle que

$$\forall k \in \mathbb{N}, \qquad \int_{-\infty}^{\infty} f_k(x)\widetilde{\nu}_n(dx) \overset{n\to\infty}{\longrightarrow} \alpha_k.$$

On définit ainsi une application ϕ de $(f_k : k \in \mathbb{N})$ dans $\mathbb{R}$. Or cette application est uniformement continue puisque pour tous $k, k' \in \mathbb{N}$,

$$|\int_{-\infty}^{\infty} f_k(x)\widetilde{\nu}(dx) - \int_{-\infty}^{\infty} f_{k'}(x)\widetilde{\nu}(dx)| \leq \| f_k - f_{k'} \|_\infty . \sup_{n\in\mathbb{N}} \widetilde{\nu}_n(I)$$

où I est un intervalle contenant le support de f_k et de $f_{k'}$. De plus $(f_k : k \in \mathbb{N})$ est dense dans $\mathcal{C}_K(\mathbb{R})$ qui est complet donc ϕ se prolonge en une application uniformement continue de $\mathcal{C}_K(\mathbb{R})$ dans $\mathbb{R}$.
Cette application est positive et le théorème de Riesz implique qu'il existe une mesure positive ν telle que pour tout $f \in \mathcal{C}_K(\mathbb{R})$, $\phi(f) = \int_{-\infty}^{\infty} f(x)\nu(dx)$.
On conclut en montrant que pour tout $f \in \mathcal{C}_K(\mathbb{R})$,

$$\begin{aligned} &|\int_{-\infty}^{\infty} f(x)\nu_n(dx) - \int_{-\infty}^{\infty} f(x)\nu_n(dx)| \\ &\leq \sup_{n\in\mathbb{N}} \nu_n(\text{supp} f \cup \text{supp} f_k) \| f - f_k \|_\infty + |\int_{-\infty}^{\infty} f_k(x)\widetilde{\nu}_n(dx) - \int_{-\infty}^{\infty} f_k(x)\nu(dx)|. \end{aligned}$$

Le terme de gauche tend vers zéro quand $n \to \infty$ en utilisant une sous suite de f_k qui converge uniformement vers f et qui est à support dans un segment $[-N,N]$ qui contient le support de f. Ceci achève la preuve.

3.1.b) Commencons par montrer que la convergence a lieu pour $f = \mathbb{1}([a,b])$ la foncton indicatrice du segment $[a,b]$. Pour cela, encadrons f par deux fonctions continues f_k^1, f_k^2 dans $[0,1]$ qui valent 0 sur $]-\infty,a] \cup [b,\infty[$ et 1 sur $[a+1/n, b-1/n]$. Alors

$$\begin{aligned} & \limsup_{n\to\infty} \int_{-\infty}^{\infty} f(x)\nu_n(dx) - \int_{-\infty}^{\infty} f(x)\nu_n(dx) \\ \leq\ & \limsup_{n\to\infty} \int_{-\infty}^{\infty} f_k^2(x)\nu_n(dx) - \int_{-\infty}^{\infty} f_k^1(x)\nu(dx) \\ \leq\ & \int_{-\infty}^{\infty} f_k^2(x)\nu(dx) - \int_{-\infty}^{\infty} f_k^1(x)\nu(dx) \\ \leq\ & \nu(]a, a+1/k[\cup]b-1/k, b[). \end{aligned}$$

qui tend vers 0 quand $k \to \infty$.
On obtient alors que la convergence a lieu pour les fonctions de la forme $f = \sum_{i=1}^{k} \lambda_k \mathbb{1}([a_i, b_i])$ puis pour les fonctions f directement Riemann intégrable sur $\mathbb{R}$ à nouveau par encadrement puisque que les petites et grandes sommes de Darboux $d_n(f), D_n(f)$ convergent vers la même valeur : $\int_{-\infty}^{\infty} f(x)dx$.

3.2)Pour tout $a, b \in \mathbb{R}$, on construit f_n une suite de fonctions à support dans $[a, b]$ tel que pour tout $x \in]a, b[$, $f_n(x)$ croit vers 1. Ceci se fair par exemple en posant f_n la fonction affine par morceaux qui vaut 0 sur $]-\infty, a] \cup [b, \infty[$ et 1 sur $[a+1/n, b-1/n]$. Alors en faisant tendre $n \to \infty$ dans $\int_{-\infty}^{\infty} f_n \nu(dx) = \int_{-\infty}^{\infty} f_n \nu(z + dx)$, la convergence monotone implique que $\nu(]a, b[) = \nu(]z + a, z + b[)$, puis $\nu([a, b[) = \nu([z + a, z + b[)$.
On peut maintenant prouver que ν est proportionelle à la mesure de Lebesgue. Posons $C = \nu([0, 1[)$. Pour tout $k, n \in \mathbb{N}$, $\nu([k/n, (k+1)/n[) = \nu([0, 1/n[)$ et par sommation on obtient $\nu([k/n, (k+1)/n[) = C/n$. Enfin pour tout $a < b$, en écrivant $[a, b[$ comme limite croissante d'ensemble de la forme $\cup_{k=k_1}^{k_2} [k/n, (k+1)/n[$, on en déduit $\nu([a, b[) = C(b - a)$.

3.3) Par 2.1), pour tout intervalle borné I, $U(I + t)$ est borné pour $t \in \mathbb{R}$. On sait que $U(t + dx)$ admet une valeur d'adhérence quand $t \to \infty$ d'après 3.1.a). Pour obtenir la convergence faible de $U(t + dx)$ quand $t \to \infty$, il suffit de prouver que l'unique valeur d'adhérence possible de cette suite est dx/m.
Soit donc une suite $t_k \to \infty$ telle que $U(t_k + dx)$ converge faiblement vers $\nu(dx)$.
Pour tout g à support compact, pour tout $a, k \geq 0$

$$f(t_k + a) = \int_{-\infty}^{\infty} g(-a)U(dx + t_k + a) \overset{k\to\infty}{\longrightarrow} \int_{-\infty}^{\infty} g(-y)\nu(a + dx).$$

D'après 2.5.c), $f(t_k + a) - f(t_k) \to 0$ quand $k \to \infty$ et on obtient $\int_{-\infty}^{\infty} g(-y)\nu(a + dx) = \int_{-\infty}^{\infty} g(-y)\nu(dx)$. D'après 3.3), il existe donc $c > 0$ tel que $\nu(dx) = cdx$.
Il reste à prouver que $c = 1/m$. On considère pour cela $g(x) = \lambda([x, \infty[)$ pour $x \geq 0$ et $g(x) = 0$ pour $x < 0$. On constate que $f(x) = g \star U(x) = 1$ pour $x \geq 0$. De plus on peut appliquer 3.1.b) puisque g est monotone d'intégrale finie sur $\mathbb{R}$ et donc directement intégrable sur $\mathbb{R}$. On obtient que $f(t_k + x) \to c \int_{-\infty}^{\infty} g(-y)dy = cm$ quand $x \to \infty$.

4.2) En notant $S_n = \sum_{k=1}^{n} X_k$, on a $S_{N_x} \leq x < S_{N_x+1}$. Or par la loi des grands nombres, $S_n/n \to m$ quand $n \to \infty$ et $N_x \to \infty$ quand $x \to \infty$. Donc $S_{N_x}/N_x \to m$ et $S_{N_x+1}/N_x \to m$ quand $x \to \infty$ p.s. Le précédent encadrement implique que N_x/x converge p.s. vers $1/m$.

5.1) On note $T_n := \sum_{i=1}^{n} X_i$ le temps de passage du $n$ième bus à la station. Alors

$$\begin{aligned}
\mathbb{P}(A_t \geq x) &= \sum_{k=0}^{\infty} \mathbb{P}(T_k < t,\ T_{k+1} \geq t + x) = \sum_{k=0}^{\infty} \mathbb{P}(T_k < t,\ X_{k+1} \geq t + x - T_k) \\
&= \sum_{k=0}^{\infty} \int_0^t \mathbb{P}(X_k \in dy)\mathbb{P}(X \geq t + x - y) = \int_0^t \mathbb{P}(X \geq t + x - y)U(dy) \\
&\to m^{-1} \int_0^t \mathbb{P}(X \geq t + x - y)dy.
\end{aligned}$$

en utilisant 3.3.a). En intégrant cette limite, on obtient la convergence de l'espérance de A_t quand $t \to \infty$.