Université des Sciences et Technologies de Lille
U.F.R. de Mathématiques Pures et Appliquées

IFP **Année 2002-03**

Corrigé du Devoir Surveillé du 27 novembre 2002

Ex 1. Soit $(\Omega, \mathcal{F}, \mathbf{P})$ un espace probabilisé et X une variable aléatoire positive définie sur cet espace.

1) Rappelons que la variable aléatoire positive X sur $(\Omega, \mathcal{F})$ n'est rien d'autre qu'une application $\Omega \to \overline{\mathbb{R}}_+$, mesurable $\mathcal{F}$-$\mathrm{Bor}(\overline{\mathbb{R}}_+)$. Par le théorème 2.23 du cours [1], X est limite (au sens de la convergence simple sur tout Ω) d'une suite croissante (X_n) de fonctions étagées mesurables positives. Ces fonctions étagées ne prennent chacune qu'un ensemble fini de valeurs distinctes. L'ensemble $X_n(\Omega)$ est donc *fini* pour tout n. Par mesurabilité de X_n pour les tribus $\mathcal{F}$ et $\mathrm{Bor}(\overline{\mathbb{R}}_+)$, $X_n^{-1}(\{x\})$ appartient à $\mathcal{F}$ pour tout $x \in \overline{\mathbb{R}}_+$, donc en particulier pour tout $x \in X_n(\Omega)$. La fonction X_n est donc bien une variable aléatoire discrète, cf. corollaire 2.12 du cours.

2) On note F_n et F les fonctions de répartition respectives de X_n et X. Pour montrer que pour tout $x \in \mathbb{R}$, $F_n(x)$ converge en décroissant vers $F(x)$, considérons dans $\mathcal{F}$ la suite des évènements $A_n := \{X_n \leq x\}$, x quelconque étant *fixé*. En raison de la croissance de la suite (X_n), cette suite (A_n) est *décroissante pour l'inclusion*. En effet si $\omega \in A_{n+1}$, alors $X_{n+1}(\omega) \leq x$ et comme $X_n(\omega) \leq X_{n+1}(\omega)$, on a aussi $X_n(\omega) \leq x$, d'où $\omega \in A_n$. Ainsi pour tout entier n, $A_{n+1} \subset A_n$.

Notons $A := \bigcap_{n\in\mathbb{N}} A_n$ et vérifions que $A = \{X \leq x\}$. D'une part si $\omega \in A$, alors ω appartient à chacun des A_n, ce qui se traduit par la suite d'inégalités $X_n(\omega) \leq x$ pour tout n. En faisant tendre n vers l'infini, on en déduit l'inégalité *large* $X(\omega) \leq x$ et donc l'appartenance de ω à A. Ainsi A est inclus dans $\{X \leq x\}$. D'autre part si $\omega \in \{X \leq x\}$, alors $X(\omega) \leq x$ et comme $X(\omega)$ est la limite dans $\overline{\mathbb{R}}_+$ de la suite croissante $\big(X_n(\omega)\big)_{n\in\mathbb{N}}$, on a pour tout n l'inégalité $X_n(\omega) \leq X(\omega)$, d'où $X_n(\omega) \leq x$ et ainsi ω appartient à *chacun* des A_n, donc aussi à leur intersection A. Ceci établit l'inclusion $\{X \leq x\} \subset A$ et achève la vérification de l'égalité de ces deux ensembles.

La *probabilité* $\mathbf{P}$ ayant la propriété de continuité séquentielle *décroissante* [2], on a

$$A_n \downarrow A \quad \Rightarrow \quad \mathbf{P}(A_n) \downarrow \mathbf{P}(A).$$

En se souvenant que $F_n(x) = \mathbf{P}(X_n \leq x) = \mathbf{P}(A_n)$ et $F(x) = \mathbf{P}(X \leq x) = \mathbf{P}(A)$, on voit que le résultat annoncé est démontré.

[1] Dans le théorème du cours, la tribu sur l'ensemble d'arrivée est $\mathrm{Bor}(\overline{\mathbb{R}})$, mais la lecture de la preuve montre immédiatement que l'on peut la remplacer par $\mathrm{Bor}(\overline{\mathbb{R}}_+)$ qui contient elle aussi tous les intervalles $[k2^{-n}, (k+1)2^{-n}[$ et $[n, +\infty]$ dont les images réciproques par X servent à construire X_n.

[2] Pour une mesure μ quelconque, cette propriété n'est valide que pour les suites décroissantes (A_n) telles que pour un certain n_0, $\mu(A_{n_0}) < +\infty$, mais cette condition est automatiquement vérifiée pour toute suite décroissante si μ est une mesure *finie*.

Ex 2. *Loi uniforme sur la boule unité en grande dimension.*
On note λ_d la mesure de Lebesgue sur $\mathbb{R}^d$, $B_d(0,r)$ la boule fermée dans $\mathbb{R}^d$ de centre 0 et de rayon r, pour la distance associée à la norme euclidienne

$$\|x\|^2 := \sum_{i=1}^d x_i^2, \quad x = (x_1, \dots, x_d).$$

et $v(d) := \lambda_d\bigl(B_d(0,1)\bigr)$.

1) Pour $r > 0$, soit $h : \mathbb{R}^d \to \mathbb{R}^d$, $x \mapsto rx$ l'homothétie de centre 0 et de rapport r. On sait que pour tout borélien A de $\mathbb{R}^d$, $\lambda_d\bigl(h(A)\bigr) = r^d\lambda_d(A)$ (cf. cours, Prop. 1.39 iv). En remarquant que $B_d(0,r) = h\bigl(B_d(0,1)\bigr)$, on en déduit immédiatement que

$$\forall r > 0, \quad \lambda_d\bigl(B_d(0,r)\bigr) = r^d v(d). \tag{1}$$

2) Soit $(\Omega, \mathcal{F}, \mathbf{P})$ un espace probabilisé et $U : \Omega \to \mathbb{R}^d$ un vecteur aléatoire sur $(\Omega, \mathcal{F})$, autrement dit une application *mesurable* $\mathcal{F}$-Bor$(\mathbb{R}^d)$. La norme euclidienne N considérée comme application $N : \mathbb{R}^d \to \mathbb{R}$, $x \mapsto \|x\|$ est continue donc borélienne, c'est-à-dire ici mesurable Bor$(\mathbb{R}^d)$-Bor$(\mathbb{R})$. On en déduit que $R := \|U\| = N \circ U$ est mesurable $\mathcal{F}$-Bor$(\mathbb{R})$ comme composée d'applications mesurables. Autrement dit R est une *variable aléatoire réelle* sur $(\Omega, \mathcal{F})$.

Supposons maintenant que la loi de U (sous $\mathbf{P}$) soit la loi uniforme sur $B_d(0,1)$, donc donnée par

$$\forall A \in \mathrm{Bor}(\mathbb{R}^d), \quad \mathbf{P}(U \in A) = \frac{\lambda_d\bigl(A \cap B_d(0,1)\bigr)}{\lambda_d\bigl(B_d(0,1)\bigr)} = \frac{1}{v(d)}\lambda_d\bigl(A \cap B_d(0,1)\bigr). \tag{2}$$

La fonction de répartition F de R est définie sur $\mathbb{R}$ par $F(r) = \mathbf{P}(R \le r)$. Si $r < 0$, $\{R \le r\} = \{\omega \in \Omega;\ \|U\| \le r\} = \emptyset$, donc $F(r) = 0$. Si $r = 0$, $\{R \le 0\} = \{R = 0\} = \{U = 0\}$ et en appliquant (2) avec $A = \{0\}$, on obtient $F(0) = 0$ puisque la mesure de Lebesgue d'un singleton dans $\mathbb{R}^d$ est nulle. Pour $r > 0$, on a l'équivalence :

$$\|U\| \le r \quad \Leftrightarrow \quad U \in B_d(0,r).$$

En appliquant alors (2) avec $A = B_d(0,r)$, on obtient :

$$\mathbf{P}(R \le r) = \mathbf{P}(\|U\| \le r) = \mathbf{P}\bigl(U \in B_d(0,r)\bigr) = \frac{1}{v(d)}\lambda_d\bigl(B_d(0,r) \cap B_d(0,1)\bigr).$$

Pour $0 < r \le 1$, $B_d(0,r) \cap B_d(0,1) = B_d(0,r)$ et pour $r > 1$, cette intersection est égale à $B_d(0,1)$. En utilisant (1), on obtient finalement :

$$F(r) = \mathbf{P}(R \le r) = \begin{cases} 0 & \text{si } r \le 0 \\ r^d & \text{si } 0 < r \le 1 \\ 1 & \text{si } r > 1. \end{cases} \tag{3}$$

Par précaution, on vérifie qu'il s'agit bien d'une fonction de répartition : elle est croissante sur $\mathbb{R}$, continue sur $\mathbb{R}$ (donc *a fortiori* continue à droite et limitée à gauche en tout point) et a pour limites 0 en $-\infty$ et 1 en $+\infty$.

Remarquons au passage que dès que $d > 1$, R ne suit pas la loi uniforme sur $[0, 1]$. On pouvait s'en douter dès le départ, en regardant le cas $d = 2$. Si U suit la loi uniforme sur le disque unité $B_2(0, 1)$, la probabilité que U « tombe » dans le disque $B_2(0, 1/10)$ est le rapport de l'aire de ce disque à celle du disque unité, soit $1/100$. La probabilité que U « tombe » dans la couronne de centre 0, de rayon intérieur $9/10$ et de rayon extérieur 1 est égale au rapport de l'aire de cette couronne à celle du disque unité, soit $1-(9/10)^2 = 1 - 81/100 = 19/100$. Ainsi $\mathbf{P}(R \in]0, 1/10]) = 1/100$, et $\mathbf{P}(R \in]9/10, 1]) = 19/100$. Si R suivait la loi uniforme sur $[0, 1]$, ces deux probabilités seraient égales toutes les deux à $1/10$.

3) Le réel m est *une* médiane de la variable aléatoire réelle R s'il vérifie à la fois $\mathbf{P}(R \leq m) \geq 1/2$ et $\mathbf{P}(R \geq m) \geq 1/2$. Il est commode d'exprimer ces deux conditions à l'aide de F. La première s'écrit simplement $F(m) \geq 1/2$. Pour la seconde, on note que

$$\mathbf{P}(R \geq m) = \mathbf{P}(R > m) + \mathbf{P}(R = m) = 1 - F(m) + \mathbf{P}(R = m).$$

Comme F est continue en tout point, la loi de R n'a pas de masse ponctuelle[3] d'où $\mathbf{P}(R = m) = 0$. Ainsi $\mathbf{P}(R \geq m) = 1-F(m)$ et la recherche des médianes de R se réduit à la résolution du systèmes d'inégalités :

$$\left\{\begin{array}{rcl} F(m) & \geq & 1/2 \\ 1 - F(m) & \geq & 1/2 \end{array}\right. \Leftrightarrow \left\{\begin{array}{rcl} F(m) & \geq & 1/2 \\ F(m) & \leq & 1/2 \end{array}\right. \Leftrightarrow \quad F(m) = 1/2.$$

Un coup d'oeil sur (3) nous assure que cette dernière équation n'a certainement aucune solution m extérieure à l'intervalle $]0, 1[$. La résolution de l'équation $F(m) = 1/2$ pour $0 < m < 1$ est immédiate puisqu'elle s'écrit $m^d = 1/2$. On a donc une solution unique $m = 2^{-1/d}$. En conclusion, R a une unique médiane :

$$\operatorname{med}(R) = 2^{-1/d}.$$

Voici une application ludique pour salle de T.P. informatique. Construisez un algorithme générant un vecteur aléatoire de loi uniforme sur la boule unité de $\mathbb{R}^3$ (par exemple par la méthode du rejet) et proposez à un autre joueur de parier que le point aléatoire généré va tomber dans la boule $B(0, r)$ tandis que vous pariez qu'il tombera dans son complémentaire dans $B(0, 1)$. Poussez le vice jusqu'à laisser votre adversaire proposer la valeur de r pour laquelle il pense que le jeu est équilibré[4]...

4) Pour $n \in \mathbb{N}^*$, notons U_n un vecteur aléatoire de loi uniforme sur $B_n(0, 1)$, $R_n := \|U_n\|$ et F_n la fonction de répartition de r_n. Pour $0 < \varepsilon < 1$, on a

$$\mathbf{P}(1 - \varepsilon < R_n \leq 1) = F_n(1) - F_n(1 - \varepsilon) = 1 - (1 - \varepsilon)^n.$$

[3]Rappelons que l'on a toujours $\mathbf{P}(R = m) = F(m) - F(m - 0)$, où $F(m - 0)$ désigne la limite à gauche de F au point m.

[4]Cette valeur est bien sûr $2^{-1/3} \simeq 0{,}794$.

Par conséquent, quand n tend vers $+\infty$, $\mathbf{P}(1-\varepsilon < R_n \leq 1)$ tend vers 1. La traduction pratique de ce résultat est qu'en grande dimension, l'essentiel de la masse de la loi uniforme sur la boule unité est concentrée au voisinage de la *sphère* unité $S_n(0,1) := \{x \in \mathbb{R}^n;\ \|x\| = 1\}$.

Pour se faire une idée, supposons que l'on ait programmé pour $n = 1000$ la simulation[5] d'un vecteur aléatoire U_n de loi uniforme sur $B_n(0,1)$. Avant de lancer ce générateur, on peut parier que le vecteur U_n aura une norme supérieure à 0,99. La probabilité de gagner ce pari est d'environ 0,999 957.

Ex 3. Le but de cet exercice est d'établir que toute f λ-intégrable sur $\mathbb{R}$, vérifie pour λ-presque tout $x \in \mathbb{R}$, l'estimation asymptotique $f(nx) = o(n^\alpha)$ quand $n \to +\infty$.

1) Commençons par examiner l'effet d'un changement d'échelle $\varphi : x \mapsto cx$ sur une intégrale de Lebesgue. L'application φ est continue, donc borélienne. Elle est aussi bijective d'inverse *ponctuel* $\varphi^{-1} : y \mapsto \frac{1}{c}y$. On en déduit facilement la détermination de l'inverse *ensembliste* d'un borélien B de $\mathbb{R}$:

$$\varphi^{-1}(B) = \{x \in \mathbb{R};\ \varphi(x) \in B\} = \{\varphi^{-1}(y);\ y \in B\} = \Big\{\frac{1}{c}y;\ y \in B\Big\} =: \frac{1}{c}B. \qquad (4)$$

Par transfert φ de l'espace mesuré $\big(\mathbb{R}, \mathrm{Bor}(\mathbb{R}), \lambda\big)$ vers $\big(\mathbb{R}, \mathrm{Bor}(\mathbb{R}), \nu := \lambda \circ \varphi^{-1}\big)$, on a pour toute $g : \mathbb{R} \to \overline{\mathbb{R}}_+$, borélienne

$$\int_{\mathbb{R}} (g \circ \varphi)\, \mathrm{d}\lambda = \int_{\mathbb{R}} g\, \mathrm{d}\nu.$$

On identifie la mesure image ν grâce à (4) et à la formule de calcul de la mesure de Lebesgue de l'homothétique d'un borélien (cf. cours, Prop. 1.39 iv) :

$$\forall B \in \mathrm{Bor}(\mathbb{R}), \quad \nu(B) = \lambda\big(\varphi^{-1}(B)\big) = \lambda\Big(\frac{1}{c}B\Big) = \frac{1}{c}\lambda(B).$$

Ceci montre que ν est la mesure $\frac{1}{c}\lambda$. On en déduit que :

$$\forall c > 0, \quad \int_{\mathbb{R}} g(cx)\, \mathrm{d}\lambda(x) = \int_{\mathbb{R}} g\, \mathrm{d}\Big(\frac{1}{c}\lambda\Big) = \frac{1}{c}\int_{\mathbb{R}} g(y)\, \mathrm{d}\lambda(y). \qquad (5)$$

2) Soit $f : \mathbb{R} \to \overline{\mathbb{R}}$, λ-intégrable sur $\mathbb{R}$ et $\alpha > 0$ une constante. En appliquant la formule de changement de variable (5) à la fonction borélienne positive $g = |f|$, on obtient

$$\sum_{n=1}^{+\infty} \int_{\mathbb{R}} |n^{-\alpha} f(nx)|\, \mathrm{d}\lambda(x) = \sum_{n=1}^{+\infty} \Big\{ \frac{1}{n^{\alpha+1}} \int_{\mathbb{R}} |f(y)|\, \mathrm{d}\lambda(y) \Big\} = \Big\{ \int_{\mathbb{R}} |f|\, \mathrm{d}\lambda \Big\} \sum_{n=1}^{+\infty} \frac{1}{n^{\alpha+1}} < +\infty.$$

En effet, l'intégrale $\int_{\mathbb{R}} |f|\, \mathrm{d}\lambda$ est finie par l'hypothèse d'intégrabilité de f et la série de Riemann de terme général $n^{-(\alpha+1)}$ converge car α est strictement positif, donc $\alpha + 1 > 1$.

[5] La méthode du rejet serait calamiteuse ici car le rapport des volumes de $[-1,1]^n$ et $B_n(0,1)$ est équivalent à $\sqrt{n\pi}\big(\frac{2n}{e\pi}\big)^{n/2}$. Il y a un algorithme bien plus performant ne nécessitant que la simulation de 1000 variables aléatoires gaussiennes $N(0,1)$ et une uniforme sur $[0,1]$.

3) Posons pour tout $x \in \mathbb{R}$, $f_n(x) := n^{-\alpha} f(nx)$. Il est clair que chaque f_n hérite de la mesurabilité de f (par composition de f avec l'homothétie continue $x \mapsto nx$ et multiplication par la constante $n^{-\alpha}$). D'après la question précédente, on a la convergence

$$\sum_{n=1}^{+\infty} \int_{\mathbb{R}} |f_n| \, \mathrm{d}\lambda < +\infty.$$

On en déduit que la série de fonctions mesurables positives $\sum_{n\geq 1} |f_n|$ est λ-intégrable sur $\mathbb{R}$, donc aussi finie λ-presque partout (cf. cours prop. 4.19 ou cor. 4.23). Par conséquent

$$\sum_{n\geq 1} |f_n(x)| < +\infty, \quad \text{pour } \lambda\text{-presque tout } x \in \mathbb{R}.$$

Comme le terme général d'une série convergente (dans $\mathbb{R}$) tend nécessairement vers zéro, on en déduit que

$$\lim_{n\to+\infty} n^{-\alpha} f(nx) = 0, \quad \text{pour } \lambda\text{-presque tout } x \in \mathbb{R}. \tag{6}$$

4) Pour voir que la clause « λ-presque tout x » n'est pas superflue dans (6), construisons un exemple de fonction f λ-intégrable sur $\mathbb{R}$ telle que : $\lim_{n\to+\infty} n^{-\alpha} f(n) = +\infty$. Ainsi la convergence dans (6) n'aura pas lieu pour $x = 1$. On en déduit facilement qu'elle n'aura alors lieu pour *aucun* $x \in \mathbb{N}^*$.

L'exemple le plus simple est sans doute $f := (+\infty)\mathbf{1}_{\mathbb{N}^*}$. Cette fonction est mesurable et nulle λ-p.p., donc λ intégrable sur $\mathbb{R}$. Pour tout $n \in \mathbb{N}^*$, $n^{-\alpha} f(n) = n^{-\alpha} \times (+\infty) = +\infty$. Un autre exemple presque aussi simple où f ne prend que des valeurs finies est $f := \sum_{k\in\mathbb{N}^*} k^{\alpha+1} \mathbf{1}_{\{k\}}$. Là encore f est mesurable et nulle λ-p.p., donc λ-intégrable. Pour tout $n \in \mathbb{N}^*$, $n^{-\alpha} f(n) = n^{-\alpha} n^{\alpha+1} = n$, donc $n^{-\alpha} f(n)$ tend vers l'infini avec n.

Dans les deux exemples précédents, (6) est évidente directement. En effet, ces deux fonctions sont nulles en dehors de $\mathbb{N}^*$. Si x n'est pas rationnel, nx ne peut appartenir à $\mathbb{N}^*$ et donc $f(nx) = 0$. Ainsi, au moins pour $x \in \mathbb{R} \setminus \mathbb{Q}$, la suite de terme général $n^{-\alpha} f(nx)$ est identiquement nulle donc trivialement convergente vers 0.

La construction suggérée par l'énoncé permet d'exhiber un exemple de fonction f avec $\int_{\mathbb{R}} f \, \mathrm{d}\lambda \neq 0$, où (6) est loin d'être évidente directement et où la convergence dans (6) n'a lieu pour aucun $x \in \mathbb{N}^*$. Cherchons donc f de la forme $f = \sum_{k=1}^{+\infty} a_k \mathbf{1}_{I_k}$, où les I_k sont des intervalles. En prenant les I_k deux à deux disjoints et en imposant à chaque I_k de contenir l'entier k, on aura pour tout $n \in \mathbb{N}^*$, $f(n) = a_n$. En prenant par exemple $a_n := n^{\alpha+1}$, on obtient bien la divergence cherchée puisque $n^{-\alpha} f(n) = n$. Il reste à vérifier que l'on peut choisir les I_k de façon que f soit effectivement λ-intégrable. La mesurabilité de f est évidente puisqu'il s'agit d'une série de fonctions mesurables positives. L'interversion série intégrale pour les fonctions mesurables positives (corollaire du th. de Beppo Levi, cf. cor. 3.12 du cours) nous donne

$$\int_{\mathbb{R}} f \, \mathrm{d}\lambda = \sum_{k=1}^{+\infty} a_k \int_{\mathbb{R}} \mathbf{1}_{I_k} \lambda = \sum_{k=1}^{+\infty} a_k \lambda(I_k).$$

En prenant par exemple $I_k = [k,\, k+\varepsilon_k]$, il suffit donc d'ajuster les longueurs ε_k des intervalles I_k de façon que la série de terme général $a_k\varepsilon_k$ soit convergente et donc que f soit intégrable. Finalement on voit que la fonction

$$f := \sum_{k=1}^{+\infty} k^{\alpha+1}\mathbf{1}_{[k,k+k^{-\alpha-3}]}$$

convient.
Remarque. La propriété (6) appliquée à cette fonction f implique que pour λ-presque tout x de $]0,1[$, la condition $nx \in I_n$ n'est vérifiée que pour un nombre *fini* d'indices n. On sait d'autre part que pour λ-presque tout x de $]0,1[$, la suite de terme général $nx - [nx]$ (*i.e.*, nx modulo 1) est *dense* dans $]0,1[$. Un peu de méditation s'impose ici pour se convaincre de la compatibilité ces deux affirmations...et de la profondeur de (6).

Problème.

On note λ la mesure de Lebesgue sur $\mathbb{R}$. Le but du problème est d'établir une inégalité entre l'intégrale de la dérivée presque partout d'une fonction croissante et la variation de cette fonction.

1) Soit $g : \mathbb{R} \to \mathbb{R}$, borélienne et λ-intégrable sur tout intervalle borné de $\mathbb{R}$. On suppose que g a une limite à droite au point a. Cela signifie qu'il existe un réel ℓ (que nous noterons ultérieurement $g(a+0)$) tel que :

$$\forall \varepsilon > 0,\ \exists \delta > 0,\ \forall x \in]a, a+\delta[, \quad \ell - \varepsilon < g(x) < \ell + \varepsilon. \tag{7}$$

Fixons $\varepsilon > 0$. On déduit de (7) que pour $0 < h < \delta$, et pour tout $x \in]a, a+h[$, $\ell - \varepsilon < g(x) < \ell + \varepsilon$. Cette inégalité montre que la fonction borélienne g est bornée sur l'intervalle de mesure finie $]a, a+h[$, donc λ-intégrable sur cet intervalle[6]. Par croissance de l'intégrale, on déduit de cet encadrement que

$$\forall h \in]0, \delta[, \qquad \int_{]a,a+h[} (\ell - \varepsilon)\,\mathrm{d}\lambda \leq \int_{]a,a+h[} g\,\mathrm{d}\lambda \leq \int_{]a,a+h[} (\ell + \varepsilon)\,\mathrm{d}\lambda. \tag{8}$$

On peut remplacer dans (8) l'intervalle d'intégration $]a, a+h[$ par $[a, a+h]$ sans changer la valeur des intégrales car la mesure de Lebesgue ne charge pas les singletons. D'autre part si c est une constante, $\int_{[a,a+h]} c\,\mathrm{d}\lambda = c\lambda([a, a+h]) = ch$. En appliquant ceci avec $c = \ell - \varepsilon$ et $c = \ell + \varepsilon$, on obtient après division par $h > 0$ dans (8) :

$$\forall h \in]0, \delta[, \quad \ell - \varepsilon \leq \frac{1}{h}\int_{[a,a+h]} g\,\mathrm{d}\lambda \leq \ell + \varepsilon. \tag{9}$$

Pour obtenir cet encadrement, nous avons travaillé avec $\varepsilon > 0$ fixé et un $\delta > 0$ dépendant de ε, donné par (7). Comme nous n'avons utilisé aucune hypothèse particulière sur ε, autre que $\varepsilon > 0$, nous avons en fait montré que :

$$\forall \varepsilon > 0,\ \exists \delta > 0,\ \forall h \in]0, \delta[, \quad \ell - \varepsilon \leq \frac{1}{h}\int_{[a,a+h]} g\,\mathrm{d}\lambda \leq \ell + \varepsilon. \tag{10}$$

[6] Donc l'hypothèse de l'énoncé « g λ-intégrable sur tout intervalle borné » est superflue.

Ceci nous donne la conclusion attendue puisque (10) est exactement l'écriture avec quantificateurs de la convergence :

$$\lim_{\substack{h\to 0\\ h>0}} \frac{1}{h}\int_{[a,a+h]} g(x)\,\mathrm{d}\lambda(x) = \ell = g(a+0). \tag{11}$$

2) Soit $F:\mathbb{R}\to\mathbb{R}$ croissante, montrons qu'elle est borélienne. Pour cela il suffit de prouver que $F^{-1}(C)$ est un borélien de $\mathbb{R}$ pour tout $C\in\mathcal{C}$, où $\mathcal{C}$ désigne une famille de boréliens telle que $\sigma(\mathcal{C})=\mathrm{Bor}(\mathbb{R})$. Prenons $\mathcal{C}:=\{]-\infty,b];\ b\in\mathbb{R}\}$. Nous allons vérifier que pour tout $b\in\mathbb{R}$, $B:=F^{-1}(]-\infty,b])$ est soit l'ensemble vide, soit un *intervalle* de $\mathbb{R}$, donc toujours un borélien de $\mathbb{R}$. En effet, si B n'est pas vide, posons

$$a:=\sup B=\sup\{t\in\mathbb{R};\ F(t)\le b\}.$$

Si $x<a$, il existe[7] $t\in B$ (c'est-à-dire vérifiant $F(t)\le b$) tel que $x<t\le a$. Or F est croissante, donc $F(x)\le F(t)\le b$, d'où $x\in B$. Ce raisonnement étant valable pour tout $x<a$, on en déduit que B contient l'intervalle $]-\infty,a[$. Si $a=+\infty$, nécessairement $B=\mathbb{R}$. Si $a<+\infty$, pour tout $x>a=\sup B$, $x\notin B$. Par conséquent B ne peut être alors que l'un des deux intervalles $]-\infty,a[$ ou $]-\infty,a]$. Ainsi dans tous les cas B est soit vide soit un intervalle de $\mathbb{R}$.

3) Si $F:\mathbb{R}\to\mathbb{R}$ est croissante, elle est λ-intégrable sur tout intervalle d'extrémités a, b, $-\infty<a<b<+\infty$. En effet, F est borélienne d'après la question précédente, il suffit donc de vérifier que $\int_{[a,b]}|F|\,\mathrm{d}\lambda<+\infty$. Si c'est vrai pour l'intervalle $[a,b]$ fermé, cela vaudra aussi pour chacun des trois autres intervalles d'extrémités a et b, puisque λ ne chargeant pas les singletons, les intégrales sur ces intervalles seront égales à $\int_{[a,b]}|F|\,\mathrm{d}\lambda$.

Par croissance de F, on a

$$\forall x\in[a,b],\quad F(a)\le F(x)\le F(b).$$

Remarquons que comme F est à valeurs dans $\mathbb{R}$, $F(a)$ et $F(b)$ sont des *réels*, donc *finis*. De l'encadrement précédent on déduit :

$$\forall x\in[a,b],\quad -F(b)\le -F(x)\le -F(a),$$

d'où[8]

$$|F(x)|=\max\bigl(F(x),-F(x)\bigr)\le\max\bigl(F(b),-F(a)\bigr)\le\max\bigl(|F(a)|,|F(b)|\bigr)=:M<+\infty.$$

Il en résulte la majoration :

$$\int_{[a,b]}|F|\,\mathrm{d}\lambda\le\int_{[a,b]}M\,\mathrm{d}\lambda=M\lambda([a,b])=M(b-a)<+\infty.$$

[7]Rappelons que le sup de B est le *plus petit* des majorants de l'ensemble B. Donc si $x<a$, x n'est pas un majorant de B et il existe au moins un élément de B strictement supérieur à x

[8]L'auteur du corrigé présente ses excuses pour cet étalage de détails sordides au lecteur qui aura spontanément pensé à l'inégalité $|F(x)|\le\max\bigl(|F(a)|,|F(b)|\bigr)$. Il précise pour sa décharge qu'il a malheureusement lu dans 90% des copies corrigées l'inégalité $|F(x)|\le|F(b)|$ et suggère aux tenants de cette inégalité de considérer l'exemple $a=-1$, $b=0$ et $F(x)=x\dots$

On suppose dans toute la suite que $F : \mathbb{R} \to \mathbb{R}$ est croissante, continue à droite en tout point de $\mathbb{R}$, dérivable λ-presque partout sur $\mathbb{R}$. On admet que l'ensemble D des points où F est dérivable est un borélien de $\mathbb{R}$ et que $\lambda(D^c) = 0$. On définit la fonction f par :

$$f(x) := \begin{cases} F'(x) & \text{si } x \in D, \\ 0 & \text{si } x \in D^c. \end{cases}$$

On se propose de prouver que pour tous réels a, b ($-\infty < a < b < +\infty$),

$$\int_{[a,b]} f \,\mathrm{d}\lambda \leq F(b) - F(a). \tag{12}$$

4) Fixons une suite (h_n) de réels strictement positifs, convergente vers 0 et posons

$$\forall x \in \mathbb{R}, \quad G_n(x) := \frac{F(x+h_n) - F(x)}{h_n}, \quad G(x) := \liminf_{n\to+\infty} G_n(x).$$

Notons que chaque G_n est définie sur $\mathbb{R}$ et à valeurs dans $\mathbb{R}_+$ à cause de la croissance de F. La fonction G est *définie sur tout* $\mathbb{R}$ (ce qui n'est pas forcément le cas pour F') et à valeurs dans $\overline{\mathbb{R}}_+$. On a

$$\forall x \in D, \quad f(x) = F'(x) = \liminf_{n\to+\infty} \frac{F(x+h_n) - F(x)}{h_n} = G(x). \tag{13}$$

Remarquons au passage que pour $x \in D$, $G(x)$ est un *réel* (donc fini). En rappelant la convention d'arithmétique dans $\overline{\mathbb{R}}_+$ « $(+\infty) \times 0 = 0$ », on déduit de (13) que f peut s'écrire

$$f = \bigl(\liminf_{n\to+\infty} G_n\bigr)\mathbf{1}_D = G\mathbf{1}_D. \tag{14}$$

Dans ce produit, le second facteur est une fonction borélienne comme indicatrice d'un borélien. C'est donc une fonction mesurable $\mathrm{Bor}(\mathbb{R})$-$\mathrm{Bor}(\mathbb{R}_+)$. Comme cette fonction ne prend que des valeurs réelles, elle reste mesurable $\mathrm{Bor}(\mathbb{R})$-$\mathrm{Bor}(\overline{\mathbb{R}}_+)$ (voir corollaire 2.11 du cours). Pour établir la mesurabilité $\mathrm{Bor}(\mathbb{R})$-$\mathrm{Bor}(\overline{\mathbb{R}}_+)$ de G, il suffit de vérifier la même mesurabilité pour chaque G_n.

La fonction F est mesurable $\mathrm{Bor}(\mathbb{R})$-$\mathrm{Bor}(\mathbb{R})$ d'après la question 2. Il en va de même pour $F(\,.\,+h_n)$, composée de F avec la translation $x \mapsto x + h_n$ qui est continue donc borélienne. On en déduit que la fonction $G_n = h_n^{-1}\bigl(F(\,.\,+h_n) - F\bigr)$ est elle aussi mesurable $\mathrm{Bor}(\mathbb{R})$-$\mathrm{Bor}(\mathbb{R})$. De plus G_n ne prenant que des valeurs positives est aussi mesurable $\mathrm{Bor}(\mathbb{R})$-$\mathrm{Bor}(\mathbb{R}_+)$, car pour tout $y \geq 0$, $G_n^{-1}([0,y]) \in \mathrm{Bor}(\mathbb{R})$ et ceci suffit pour avoir la mesurabilité de G_n pour $\mathrm{Bor}(\mathbb{R}_+)$, cf. corollaire 2.10 iv) du cours. Finalement G_n est aussi mesurable $\mathrm{Bor}(\mathbb{R})$-$\mathrm{Bor}(\overline{\mathbb{R}}_+)$ par une nouvelle invocation du corollaire 2.11 et cette mesurabilité passe à $G = \liminf G_n$. Revenant à (14), on conclut que f est mesurable $\mathrm{Bor}(\mathbb{R})$-$\mathrm{Bor}(\overline{\mathbb{R}}_+)$ comme produit de deux fonctions ayant cette même mesurabilité. De plus comme f ne prend que des valeurs finies, elle est aussi mesurable $\mathrm{Bor}(\mathbb{R})$-$\mathrm{Bor}(\mathbb{R}_+)$.

5) Les fonctions f et G sont toutes deux mesurables Bor($\mathbb{R}$)-Bor($\overline{\mathbb{R}}_+$), donc l'ensemble $A := \{f \neq G\}$ est un borélien de $\mathbb{R}$. D'après la définition de f on a clairement $A \subset D^c$ d'où $\lambda(A) = 0$. Ainsi les fonctions boréliennes positives f et G sont égales λ-presque partout sur $\mathbb{R}$, elles ont donc même intégrale sur $[a,b]$, ce qui s'écrit :

$$\int_{[a,b]} f(x)\,\mathrm{d}\lambda(x) = \int_{[a,b]} \liminf_{n\to+\infty} \frac{F(x+h_n)-F(x)}{h_n}\,\mathrm{d}\lambda(x). \tag{15}$$

6) Soit $\varphi : \mathbb{R} \to \mathbb{R}$ la translation $x \mapsto x+h$. En raison de l'invariance de la mesure de Lebesgue par translation, on a $\lambda \circ \varphi^{-1} = \lambda$. Par transfert φ de l'espace mesuré $\big(\mathbb{R}, \mathrm{Bor}(\mathbb{R}), \lambda\big)$ vers $\big(\mathbb{R}, \mathrm{Bor}(\mathbb{R}), \nu := \lambda\circ\varphi^{-1} = \lambda\big)$, on a pour toute $g : \mathbb{R} \to \mathbb{R}$, ν-intégrable sur $\mathbb{R}$,

$$\int_{\mathbb{R}} (g\circ\varphi)\,\mathrm{d}\lambda = \int_{\mathbb{R}} g\,\mathrm{d}\nu = \int_{\mathbb{R}} g\,\mathrm{d}\lambda. \tag{16}$$

En raison de l'équivalence entre les encadrements $a \leq x \leq b$ et $a+h \leq x+h \leq b+h$, on a l'égalité $\mathbf{1}_{[a,b]}(x) = \mathbf{1}_{[a+h,b+h]}(x+h)$, d'où

$$\int_{[a,b]} F(x+h)\,\mathrm{d}\lambda(x) = \int_{\mathbb{R}} F(x+h)\mathbf{1}_{[a,b]}(x)\,\mathrm{d}\lambda(x) = \int_{\mathbb{R}} F(x+h)\mathbf{1}_{[a+h,b+h]}(x+h)\,\mathrm{d}\lambda(x).$$

Cette dernière intégrale est de la forme $\int_{\mathbb{R}}(g\circ\varphi)\,\mathrm{d}\lambda$ avec $g = F\mathbf{1}_{[a+h,b+h]}$ et g est bien λ-intégrable sur $\mathbb{R}$ puisque F est λ-intégrable sur tout intervalle d'extrémités finies (ici $a+h$ et $b+h$) par la question 3. La formule de transfert (16) nous donne

$$\int_{\mathbb{R}} F(x+h)\mathbf{1}_{[a+h,b+h]}(x+h)\,\mathrm{d}\lambda(x) = \int_{\mathbb{R}} F(y)\mathbf{1}_{[a+h,b+h]}(y)\,\mathrm{d}\lambda(y) = \int_{[a+h,b+h]} F\,\mathrm{d}\lambda.$$

Nous venons ainsi d'établir pour tout réel h, l'égalité :

$$\int_{[a,b]} F(x+h)\,\mathrm{d}\lambda(x) = \int_{[a+h,b+h]} F(y)\,\mathrm{d}\lambda(y). \tag{17}$$

Comme $h_n > 0$ tend vers zéro, il existe un entier n_0 tel que pour tout $n \geq n_0$, $a < a+h_n < b$. On a alors successivement (voir justifications ci-dessous) :

$$\begin{aligned}
\int_{[a,b]} \big(F(x+h_n)-F(x)\big)\,\mathrm{d}\lambda(x) &= \int_{[a,b]} F(x+h_n)\,\mathrm{d}\lambda(x) - \int_{[a,b]} F(x)\,\mathrm{d}\lambda(x) && (18)\\
&= \int_{[a+h_n,b+h_n]} F\,\mathrm{d}\lambda - \int_{[a,b]} F\,\mathrm{d}\lambda && (19)\\
&= \int_{[a+h_n,b[} F\,\mathrm{d}\lambda + \int_{[b,b+h_n]} F\,\mathrm{d}\lambda \\
&\quad - \int_{[a,a+h_n]} F\,\mathrm{d}\lambda - \int_{]a+h_n,b]} F\,\mathrm{d}\lambda && (20)\\
&= \int_{[b,b+h_n]} F\,\mathrm{d}\lambda - \int_{[a,a+h_n]} F\,\mathrm{d}\lambda. && (21)
\end{aligned}$$

Justifications :

(18) : par linéarité de l'intégrale.

(19) : par application de (17) avec $h = h_n$.

(20) : si A et B sont deux boréliens *disjoints* et F est intégrable sur $A \cup B$, elle l'est aussi sur A et sur B et $\int_{A\cup B} F \,\mathrm{d}\lambda = \int_A F \,\mathrm{d}\lambda + \int_B F \,\mathrm{d}\lambda$. C'est une conséquence immédiate de la définition de l'intégrabilité et de la relation $\mathbf{1}_{A\cup B} = \mathbf{1}_A + \mathbf{1}_B$. On applique ceci avec $A = [a + h_n, b[$ et $B = [b, b + h_n]$ d'une part et avec $A' = [a, a + h_n]$, $B' =]a + h_n, b]$ d'autre part. Signalons à cette occasion qu'il n'y a pas à proprement parler de relation de Chasles pour l'intégrale de Lebesgue. En effet cette relation dans le cadre de l'intégrale de Riemann exploite la convention $\int_c^d f(x) \,\mathrm{d}x = -\int_d^c f(x) \,\mathrm{d}x$. Si $c < d$, on n'a pas en général pour une intégrale de Lebesgue la relation $\int_{[c,d]} f \,\mathrm{d}\lambda = \int_{[d,c]} f \,\mathrm{d}\lambda$, tout simplement parce que l'intervalle $[d, c]$ est l'ensemble vide (c'est l'ensemble des $x \in \mathbb{R}$ tels que $d \leq x \leq c$), ce qui annule automatiquement la seconde intégrale.

(21) : comme la mesure de Lebesgue ne charge par les singletons, on a $\int_{[a+h_n,b[} F \,\mathrm{d}\lambda = \int_{]a+h_n,b]} F \,\mathrm{d}\lambda$, ce qui légitime la simplification effectuée pour obtenir (21).

Finalement, en utilisant (21) et la linéarité de l'intégrale, on a pour tout $n \geq n_0$:

$$\int_{[a,b]} \frac{F(x + h_n) - F(x)}{h_n} \,\mathrm{d}\lambda(x) = \frac{1}{h_n} \int_{[b,b+h_n]} F \,\mathrm{d}\lambda - \frac{1}{h_n} \int_{[a,a+h_n]} F \,\mathrm{d}\lambda.$$

En appliquant le résultat de la question 1 avec $g = F$ et en notant que les limites à droite $F(b + 0)$ et $F(a + 0)$ sont *finies*, on en déduit que

$$\lim_{n\to+\infty} \int_{[a,b]} \frac{F(x + h_n) - F(x)}{h_n} \,\mathrm{d}\lambda(x) = F(b + 0) - F(a + 0). \tag{22}$$

7) Nous pouvons maintenant conclure la preuve de (12). Il suffit de combiner (15), le lemme de Fatou appliqué à la suite de fonctions mesurables positives (G_n), (22) et la continuité à droite de F aux points a et b :

$$\begin{aligned}
\int_{[a,b]} f(x) \,\mathrm{d}\lambda(x) &= \int_{[a,b]} \liminf_{n\to+\infty} \frac{F(x + h_n) - F(x)}{h_n} \,\mathrm{d}\lambda(x) \\
&\leq \liminf_{n\to+\infty} \int_{[a,b]} \frac{F(x + h_n) - F(x)}{h_n} \,\mathrm{d}\lambda(x) \\
&= \lim_{n\to+\infty} \int_{[a,b]} \frac{F(x + h_n) - F(x)}{h_n} \,\mathrm{d}\lambda(x) \\
&= F(b + 0) - F(a + 0) = F(b) - F(a).
\end{aligned}$$

8) L'inégalité dans (12) peut être stricte, comme le montre l'exemple $F := \mathbf{1}_{[0,+\infty[}$. Cette fonction est croissante sur $\mathbb{R}$, continue à droite en tout point, dérivable partout sauf en $x = 1$ et de dérivée nulle. La fonction f associée est donc la fonction nulle sur $\mathbb{R}$, d'où $\int_{[0,1]} f \,\mathrm{d}\lambda = 0 < F(1) - F(0) = 1 - 0 = 1$.

Signalons qu'il existe aussi des exemples (beaucoup moins élémentaires) de fonctions F *continues* sur tout $\mathbb{R}$ pour laquelle l'inégalité (12) est stricte : voir le problème sur l'escalier de Cantor dans les Annales d'IFP 2001–2002.