Examen (durée 3h)
Jeudi 15 Janvier 2009
Les notes de cours ne sont pas autorisées

Les réponses aux questions doivent être rédigées avec soin.

Barème indicatif : Q.C. 2 pts, Ex1 3 pts, Ex2 5 pts, Ex 3 3pts, Problème 10 pts

Questions de cours. Enoncer la loi forte des grands nombres et le théorème de Levy.

Exercice 1. On considère un espace mesuré $(E, \mathcal{A}, \mu)$ et une suite $(f_n)_{n\geq 0}$ de densités de probabilité ($f_n \geq 0$ μ-presque partout et $\int f_n d\mu = 1$) convergeant μ-presque partout vers une densité de probabilité f.

(1) (a) Vérifier que $\int |f - f_n| d\mu = 2\int (f - f_n)\mathbb{1}_{(f-f_n)(x)\geq 0}\, d\mu(x)$.
   (b) Déduire que f_n converge vers f dans L^1.

(2) Soient $(X_n)_{n\geq 0}$ une famille de variable aléatoire de loi $f_n\mu$ pour tout $n \geq 0$ et X une variable aléatoire de loi $f\mu$, montrer que X_n converge en loi vers X.

Exercice 2. Soit $\mathcal{M}_1(\mathbb{R}^d)$ l'ensemble des mesures de probabilités boréliennes sur $\mathbb{R}^d$. Pour tous $\mu, \nu \in \mathcal{M}_1(\mathbb{R}^d)$, on définit

$$\|\mu - \nu\| \doteq 2 \sup_{A\subset \mathcal{B}(\mathbb{R}^d)} |\mu(A) - \nu(A)|\,.$$

(1) Vérifier que $\| \ \|$ satisfait l'inégalité triangulaire sur $\mathcal{M}_1(\mathbb{R}^d)$.

(2) Soient f_1 et f_2 boréliennes positives telles que $\mu = f_1(\mu+\nu)$ et $\nu = f_2(\mu+\nu)$. Montrer que

$$\|\mu - \nu\| = \int |f_1 - f_2| d(\mu+\nu)\,.$$

(3) Montrer que pour toute $f \in C_b(\mathbb{R}^d, \mathbb{R})$ (continue bornée) on a $|\int f d\mu - \int f d\nu| \leq \|\mu - \nu\| \|f\|_\infty$ et que

$$\|\mu - \nu\| = \sup_{f\in C_b(\mathbb{R}^d,\mathbb{R}), \|f\|_\infty \leq 1} |\mu(f) - \nu(f)|\,.$$

(*Pour le dernier point on pourra penser à un argument de densité des fonctions continues bornées dans un espace approprié*).

(4) Montrer que si une suite $(\mu_n)_{n\geq 0} \in \mathcal{M}_1(\mathbb{R}^d)^{\mathbb{N}}$ tend vers $\mu_\infty \in \mathcal{M}_1(\mathbb{R}^d)$ pour la norme $\| \ \|$, alors μ_n converge étroitement vers μ_∞. Que dire de la réciproque ?

Exercice 3. On considère pour tout $n \geq 1$ une suite $(X_{n,m})_{1\leq m\leq n}$ de v.a.r. indépendantes telles que

- $P(X_{n,m} = 1) = p_{n,m}$ et $P(X_{n,m} = 0) = 1 - p_{n,m}$,
- $\sum_{m=1}^n p_{n,m} \to \lambda \in]0, \infty[$ lorsque $n \to \infty$,
- $\max_{1\leq m\leq n} p_{n,m} \to 0$ lorsque $n \to \infty$.

(1) On note $S_n \doteq \sum_{m=1}^n X_{n,m}$.

(a) Calculer $\phi_n(t) \doteq E(e^{itS_n})$ et montrer qu'il existe $C > 0$ tel que

$$|\phi_n(t) - e^{\sum_{m=1}^n p_{n,m}(e^{it}-1)}| \leq C \sum_{m=1}^{n} p_{n,m}^2 .$$

(b) En déduite la limite de $\phi_n(t)$ lorsque $t \to \infty$.

(2) Calculer $E(e^{itX})$ lorsque $X \sim \mathcal{P}(\lambda)$ i.e. $P(X = k) = \lambda^k e^{-\lambda}/k!$ pour tout $k \in \mathbb{N}$ et conclure.

Problème. Soient $(\Omega, \mathcal{F}, P)$ un espace probabilisé et $(U_i)_{1\leq i\leq n}$ une famille de variables aléatoires indépendantes de loi uniforme sur $[0, 1]$. En posant $M_i \doteq (\cos(2\pi U_i), \sin(2\pi U_i))$ pour tout $1 \leq i \leq n$, on définit une famille de variables indépendantes de loi uniforme sur le cercle unité de $\mathbb{R}^2$. Les points $\{M_1, \cdots , M_n\}$ subdivisent le cercle unité en n arcs de cercle.

Pour tout $\omega \in \Omega$, on ordonne les valeurs prises par les $U_i(\omega)$ par ordre croissant et on définit $V_1(\omega) \doteq \min_{1\leq i\leq n} U_i(\omega) \leq V_2(\omega) \leq \cdots \leq V_n(\omega) \doteq \max_{1\leq i\leq n} U_i(\omega)$ les variables aléatoires associées. On numérote maintenant les n arc de cercles $(A_i)_{1\leq i\leq n}$ en partant du point $(\cos(2\pi V_1), \sin(2\pi V_1))$ et en suivant le sens direct. Ainsi, pour tout $1 \leq i \leq n - 1$, A_i est de longueur $l(A_i) \doteq 2\pi(V_{i+1} - V_i)$ et l'arc A_n désignant l'arc recouvrant le point $(1, 0)$ est de longueur $l(A_n) = 2\pi(V_1 + (1 - V_n))$ (voir Fig 1).

(1) Soit I une v.a. de loi uniforme dans $\{1, \cdots , n\}$ et indépendante des U_i. On définit $L \doteq l(A_I)$ i.e. $L = \sum_{i=1}^n l(A_i)\mathbb{1}_{I=i}$ la longueur d'un arc tiré au hasard. En remarquant que $\sum_{i=1}^n l(A_i) = 2\pi$, calculer $E(L)$.

(2) (a) Calculer $P(V_1 > x)$ pour tout $x \in [0, 1]$ et en déduire la loi de V_1.

(b) On note $L_n \doteq l(A_n)$ la longueur de l'arc chevauchant $(1, 0)$. Vérifier que V_1 et $1 - V_n$ ont même loi puis montrer que $E(L_n) = \frac{4\pi}{n+1}$.

(c) Pourquoi la longueur de l'arc A_n est en espérance plus grande que la longueur moyenne des arcs ? Essayer d'expliquer simplement ce paradoxe.

(3) Soit $(Z_i)_{i\geq 1}$ une famille de variable aléatoire i.i.d de loi exponentielle de paramètre 1 (ie $P(Z_i \geq x) = \exp(-x)$ pour tout $x \geq 0$). On note $T_k = \sum_{i=1}^k Z_i$ pour tout $k \geq 1$.

(a) Montrer que pour tout f borélienne bornée,

$$E(f(T_1, \cdots , T_{n+1})) = \int_{\mathbb{R}^{n+1}} f(t_1, \cdots , t_{n+1})\mathbb{1}_{0<t_1<\cdots<t_{n+1}} e^{-t_{n+1}} dt_1 \cdots dt_{n+1} .$$

et en déduire la loi de $(T_1, \cdots , T_{n+1})$.

(b) Montrer que $(T_1/T_{n+1}, \cdots , T_n/T_{n+1})$ et T_{n+1} sont indépendants et que $(V_1, \cdots , V_n)$ et $(T_1/T_{n+1}, \cdots , T_n/T_{n+1})$ ont même loi (*on pourra utiliser sans démonstration que la loi de* $(V_1, \cdots , V_n)$ *est absolument continue par rapport à la mesure de Lebesgue* λ_n *et de densité* $g(v_1, \cdots , v_n) = n!\mathbb{1}_{0\leq v_1\leq v_2\cdots\leq v_n\leq 1}$).

(4) (a) Montrer que $\frac{T_{n+1}}{n+1}$ admet une limite *p.s.* puis que $\frac{nZ_k}{T_{n+1}}$ converge en loi vers la loi de Z_k lorsque $n \to \infty$.

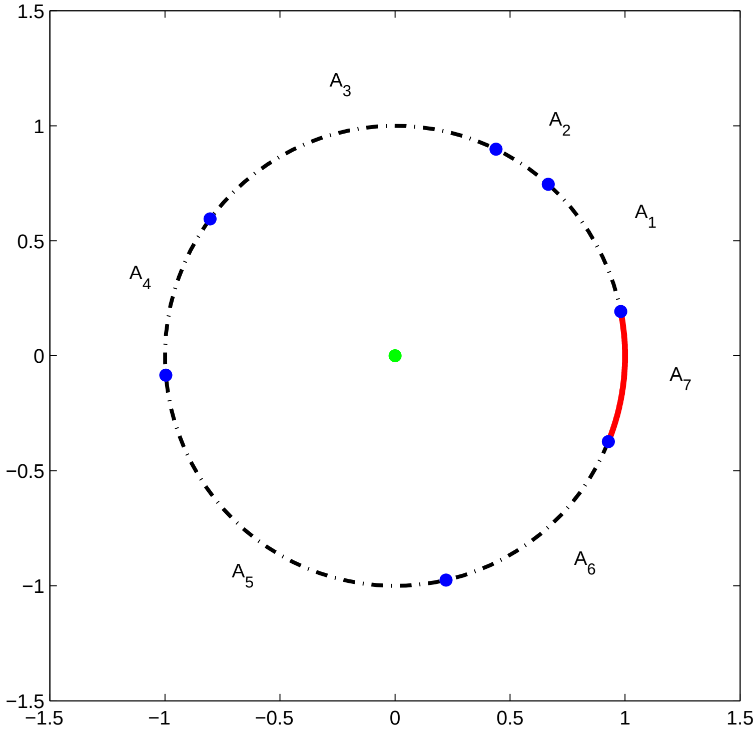

FIG. 1. Exemple de configuration pour $n = 7$. L'arc de cercle A_7, l'arc chevauchant le point $(1, 0)$ est souligné en trait plein.

(b) On pose $Y_n = \max_{1\leq k\leq n} Z_k/\log(n)$. Calculer pour tout $y \geq 0$ la valeur de $P(Y_n \leq y)$ et montrer que Y_n tend en probabilité vers 1.

(5) On note $L_* = \max_{1\leq i\leq n} l(A_i)$ la longueur du plus grand arc. Montrer alors que $nL_*/\log(n) \to 2\pi$ en probabilité avec 1.

(6) (bonus) Que dire de la longueur du plus petit arc ?

CORRECTION

Exercice1.

(1) (a) On écrit $\int |f - f_n|d\mu = \int (f - f_n)\mathbb{1}_{f-f_n\geq 0}d\mu + \int (f_n - f)\mathbb{1}_{f_n-f\geq 0}d\mu$. Or $\int (f - f_n)d\mu = 0$ et donc $\int (f_n - f)\mathbb{1}_{f_n-f\geq 0}d\mu = \int (f - f_n)\mathbb{1}_{f-f_n\geq 0}d\mu = 0$. Par suite on déduit le résultat.

(b) Comme $0 \leq (f - f_n)\mathbb{1}_{f-f_n\geq 0} \leq f$, on déduit le résultat par convergence dominée pour le chapeau intégrable f.

(2) En utilisant le théorème de Levy et comme

$$|\int e^{itx}(f_n - f)d\mu| \leq \int |f_n - f|d\mu \to 0$$

on déduit le résultat. On peut également se passer du thm de Levy en utilisant plus directement la définition de la convergence en loi : pour tout g continu borné, on a $|E(g(X_n)-g(X))| = |\int g(x)(f_n(x)-f(x))d\mu(x)| \leq |g|_\infty \int |f_n(x)-f(x)|d\mu(x) \to 0$ d'après la question précédente.

Exercice 2.

(1) Il suffit de remarquer que pour tout $A \in \mathcal{B}(\mathbb{R}^d)$, on a $|\mu(A) - \mu''(A)| \leq |\mu(A) - \mu'(A)| + |\mu'(A) - \mu''(A)|$.

(2) Pour tout $A \in \mathcal{B}(\mathbb{R}^d)$, on a $|\mu(A) - \nu(A)| = |\int_A (f_1 - f_2) d(\mu+\nu)|$. On comme $\int (f_1 - f_2) d(\mu+\nu) = 0$, on a $|\int_A (f_1 - f_2) d(\mu+\nu)| = |\int_{A^c} (f_1 - f_2) d(\mu+\nu)|$ d'où

$$|\mu(A) - \nu(A)| = \frac{1}{2}(|\int_A (f_1 - f_2) d(\mu+\nu)| + |\int_{A^c} (f_1 - f_2) d(\mu+\nu)| \leq \frac{1}{2}\int |f_1 - f_2| d(\mu+\nu)\,.$$

En prenant le sup sur A, on obtient

$$\|\mu - \nu\| \leq \int |f_1 - f_2| d(\mu+\nu),.$$

En prenant maintenant $A = \{x \in \mathbb{R}^d \mid (f_1 - f_2)(x) \geq 0\}$, on obtient facilement l'égalité.

(3) On a $|\int f d\mu - \int f d\nu| = |\int f(f_1 - f_2) d(\mu+\nu)| \leq \|f\|_\infty \int |f_1 - f_2| d(\mu+\nu)$ d'où

$$\sup_{f \in C_b(\mathbb{R}^d, \mathbb{R}), \|f\|_\infty \leq 1} |\mu(f) - \nu(f)| \leq \|\mu - \nu\|\,.$$

De plus en notant $g = \mathbb{1}_{f_1 - f_2 \geq 0} - \mathbb{1}_{f_1 - f_2 < 0}$, on a d'après 2) que $\|\mu(g) - \nu(g)\| = \int g(f_1 - f_2) d(\mu+\nu) = \int |f_1 - f_2| d(\mu+\nu)$. Par densité des fonctions continues à support compactes dans $L^1(\mu+\nu)$ il existe une suite $(g_n)_{n\geq 0}$ de fonctions continues bornées telles que $(\mu+\nu)(|g_n - g|) \to 0$. Or quitte à prendre $\tilde{g}_n \doteq -1 \vee (g_n \wedge 1)$ (pour lequel $(\mu+\nu)(|\tilde{g}_n - g|) \leq (\mu+\nu)(|g_n - g|)$ puisque $\|g\|_\infty = 1$) on peut supposer que $\|g_n\|_\infty \leq 1$. Or $|(\int g d\mu - \int g d\nu) - (\int g_n d\mu - \int g_n d\nu)| \leq \int |g - g_n| d(\mu+\nu) \to 0$. Par suite $\sup_{f \in C_b(\mathbb{R}^d, \mathbb{R}), \|f\|_\infty \leq 1} |\mu(f) - \nu(f)| \geq \lim_{n\to\infty} |\mu(g_n) - \nu(g_n)| = \|\mu - \nu\|$ d'où le résultat.

(4) Si $\|\mu_n - \mu_\infty\| \to 0$ alors pour tout $f \in C_b(\mathbb{R}^d, \mathbb{R})$, $|\mu_n(f) - \mu_\infty(f)| \leq \|f\|_\infty \|\mu_n - \mu_\infty\| \to 0$ et donc μ_n converge étroitement vers μ_∞. La réciproque est fausse, en effet, sur $\mathbb{R}$, $\delta_{1/n} \to \delta_0$ mais pour tout $n > 0$, on a $\|\delta_n - \delta_0\| = 2$.

Exercice 3.

(1) (a) On a $\phi_n(t) \doteq \prod_{m=1}^n (p_{n,m} e^{it} + (1 - p_{n,m})) = \prod_{m=1}^n (p_{n,m}(e^{it} - 1) + 1)$. Or en utilisant Taylor-Lagrange, on a pour $|e^u - (1+u)| \leq \exp(|u|)|u|^2/2$ pour $u \in \mathbb{C}$. Par suite, comme $|p_{n,m} e^{it} + (1 - p_{n,m})| \leq 1$ on déduit de la majoration

$$|\prod_{m=1}^n a_m - \prod_{m=1}^n a'_m| \leq \sum_{k=1}^n |\prod_{m<k} a_m (a_k - a'_k) \prod_{m>k} a'_m|$$

pour $a_m = p_{n,m} e^{it} + (1 - p_{n,m})$ et $a'_m = e^{p_{n,m}(e^{it}-1)}$, que

$$\begin{aligned} |\phi_n(t) - e^{\sum_{m=1}^n p_{n,m}(e^{it}-1)}| &\leq \sum_{k=1}^n \prod_{m<k} (p_{n,m} e^{it} + (1 - p_{n,m}))(p_{n,k}^2/2) \prod_{m=k}^n e^{p_{n,m}|e^{it}-1|} \\ &\leq (\sum_{k=1}^n p_{n,k}^2) e^{\sum_{m=1}^n p_{n,m}|e^{it}-1|}/2 \\ &\leq C(\sum_{k=1}^n p_{n,k}^2)\,, \end{aligned}$$

en notant $C \doteq \sup_{n\geq 0}\exp(\sum_{m=1}^n p_{n,m}|e^{it}-1|)/2<\infty$ puisque $\sum_{m=1}^n p_{n,m}\to\lambda$ lorsque $n\to\infty$.

(b) En passant à la limite, on obtient $\phi_n(t)\to e^{\lambda(e^{it}-1)}$.

(2) On calcule directement que $E(e^{itX}) = e^{-\lambda}\sum_{k\geq 0}(\lambda e^{it})^k/k! = e^{\lambda(e^{it}-1)}$ ce qui donne le résultat en utilisant le théorème de Levy.

Problème.

(1) $E(L)=\sum_{i=1}^n E(l(A_i)\mathbb{1}_{I=i})$. Or comme les variables I et $(U_i)_{1\leq i\leq n}$ sont indépendantes, $E(l(A_i)\mathbb{1}_{I=i}) = E(l(A_i))P(I=i)$ et donc $E(L)=\sum_{i=1}^n E(l(A_i))/n = E(\sum_{i=1}^n l(A_i))/n = \frac{2\pi}{n}$.

(2) (a) $P(V_1>x)=P(U_1>x)^n=(1-x)^n=\int_0^1 n(1-u)^{n-1}\mathbb{1}_{]x,1]}du$ pour $x\in[0,1]$. On déduit que la loi de V_1 est absolument continue par rapport à la mesure de Lebesgue et de densité $n(1-x)^{n-1}\mathbb{1}_{[0,1]}$.

(b) On remarque que $(U_1,\cdots,U_n)$ et $((1-U_1),\cdots,1-U_n)$ ont même loi. Or $\min(1-U_i)=1-\max U_i=1-V_n$ V_1 et $1-V_n$ ont même loi. Comme $E(L_n)=2\pi(E(V_1)+E(1-V_n))$, on déduit que $E(L_n)=4\pi E(V_1)$. On termine en remarquant que $E(V_1)=\int_0^1 P(V_1>x)dx=\frac{1}{n+1}$.

(c) Dans le cas de L_n, les arcs ont d'autant plus de chance d'être sélectionnés qu'ils sont grands ce qui augmente la valeur de l'espérance. Une autre façon d'obtenir le résultat est de considérer que l'on tire $n+1$ points et que le dernier point tiré sert à sélectionner l'intervalle (ie joue le rôle du point $(1,0)$). Dans ce cas, l'intervalle sélectionné est constitué de deux intervalles contigus pour la partition du cercle en $n+1$ intervalle d'où une longueur $2\times 2\pi/(n+1)$.

(3) (a) Il suffit de considérer le changement de variable C^∞ de $(z_1,\cdots,z_n)\to(z_1,z_1+z_2,\cdots,z_1+\cdots+z_n)$ de $((\mathbb{R}_+^*)^n)$ dans l'ouvert $\{(t_1,\cdots,t_n)\in\mathbb{R}^n \mid 0<t_1<\cdots<t_n\}$. Le Jacobien de la transformation est triangulaire de déterminant 1 si bien que

$$
\begin{aligned}
&E(f(T_1,\cdots,T_n))\\
&\overset{\text{ind. des } Z_i}{=} \int_{\mathbb{R}^n} f(z_1,z_1+z_2,\cdots,z_1+\cdots+z_n)\left(\prod_{i=1}^n e^{z_i}\mathbb{1}_{z_i>0}\right)dz_1\cdots dz_n\\
&= \int_{\mathbb{R}_+^n} f(z_1,z_1+z_2,\cdots,z_1+\cdots+z_n)e^{\sum_{i=1}^n z_i}dz_1\cdots dz_n\\
&= \int_{\mathbb{R}^{n+1}} f(t_1,\cdots,t_n)\mathbb{1}_{0<t_1<\cdots<t_n}e^{-t_n}dt_1\cdots dt_n\,. \quad (1)
\end{aligned}
$$

si bien que $(T_1,\cdots,T_n)$ admet une loi à densité sur $\mathbb{R}^n$ (par rapport à λ_n) de densité $g(t_1,\cdots,t_n)=\mathbb{1}_{0<t_1<\cdots<t_n}e^{-t_n}$. En effet si $\mu=g\lambda_n$, alors

$$E(f(T_1,\cdots,T_n)=\int f(t_1,\cdots t_n)d\mu(t_1,\cdots,t_n)$$

pour $f\in C_b(\mathbb{R}^n,\mathbb{R})$ si bien que la loi de $(T_1,\cdots,T_n)$ et μ ont même fonction caractéristique et donc sont égales par le théorème injectivité de la transformée de Fourier.

(b) On calcule en faisant le changement de variables C^1, $(t_1, \cdots, t_{n+1}) \overset{\psi}{\mapsto} (u_1 = t_1/t_{n+1}, \cdots, u_n = t_n/t_{n+1}, t_{n+1})$ d'inverse $(u_1, \cdots, u_n, t_{n+1}) \mapsto (u_1 t_{n+1}, \cdots, u_n t_{n+1}, t_{n+1})$ que pour f, g boréliennes bornées on a

$$\begin{aligned}
&E(f(T_1/T_{n+1}, \cdots, T_n/T_{n+1})g(T_{n+1})) \\
&= \int_{\mathbb{R}^{n+1}} f(t_1/t_{n+1}, \cdots, t_n/t_{n+1})g(t_{n+1})\mathbb{1}_{0<t_1<\cdots<t_{n+1}} e^{-t_{n+1}} dt_1 \cdots dt_{n+1} \\
&= \int_{]0,1[^n \times \mathbb{R}_+^*} f(u_1, \cdots, u_n)g(t_{n+1})\mathbb{1}_{0<u_1<\cdots<u_n<1} t_{n+1}^{n+1} e^{t_{n+1}} du_1 \cdots du_n dt_{n+1} \qquad (2) \\
&= \int_{]0,1[^n} f(u_1, \cdots, u_n)\left(n!\mathbb{1}_{0<u_1<\cdots<u_n<1}\right) du_1 \cdots du_n \int_{\mathbb{R}_+^*} g(t_{n+1}) \frac{1}{n!} t_{n+1}^{n+1} e^{t_{n+1}} dt_{n+1}\,.
\end{aligned}$$

En prenant $g \equiv 1$, on obtient que la loi de $(T_1/T_{n+1}, \cdots, T_n/T_{n+1})$ est celle de $(V_1, \cdots, V_n)$ et en prenant $f \equiv 1$ que celle de T_{n+1} est celle d'une loi $\gamma(n, 1)$. Enfin la dernière égalité n'est rien d'autre que

$$E(f(T_1/T_{n+1}, \cdots, T_n/T_{n+1})g(T_{n+1})) = E(f(T_1/T_{n+1}, \cdots, T_n/T_{n+1}))E(g(T_{n+1}))$$

ce qui caractérise l'indépendance.

(4) (a) La loi des grands nombres appliquées à la famille $(Z_i)_{i\geq 1}$ de v.a.i.i.d. intégrables donne $T_{n+1}/(n+1) \to E(Z_1) = 1$ presque sûrement. Notons $A_n = \frac{n}{T_{n+1}}$ et montrons que $A_n Z_n$ converge en loi vers la loi de Z_1. Il nous suffit de montrer que pour toute f continue bornée, $|E(f(A_n Z_n) - f(Z_n))| \to 0$. Or $|E(f(A_n Z_n) - f(Z_n))| \leq 2|f|_\infty P(|Z_n| \geq A) + 2|f|_\infty P(|A_n - 1| \geq \eta) + \sup_{|x|\leq A, |x-y|\leq\eta} |f(x) - f(y)|$. Pour tout $\epsilon > 0$, il existe $A > 0$ tel que $2|f|_\infty P(|Z_n| \geq A) \leq \epsilon$, puis par uniforme continuité de f sur les compacts il existe $\eta > 0$ tel que $\sup_{|x|\leq A, |x-y|\leq\eta} |f(x) - f(y)| \leq \epsilon$ et enfin par c.d. $N > 0$ tel que pour tout $n \geq N$ $2|f|_\infty P(|A_n - 1| \geq \eta) \leq \epsilon$.

(b) On a $P(Y_n \leq y) = P(\cap_{i=1}^n (Z_i \leq y \log(n))) = P(Z_1 \leq y\log(n))^n$. Or $P(Z > x) = e^{-x}$ d'où $P(Y_n \leq y) = \exp(n \log(1 - e^{-y\log(n)}))$. Comme $n\log(1 - e^{-y\log(n)}) \to -\infty$ pour $y < 1$ et vers 0 pour $y > 1$, on déduit $P(Y_n \leq y) \to 0$ pour $y < 1$ et $P(Y_n \leq y) \to 1$ pour $y > 1$. En particulier, $P(|Y_n - 1| \leq \epsilon) \to 0$ avec n, i.e. Y_n converge en probabilité vers 1.

(5) D'après 3b), $nL_*/\log(n)$ a la même loi que

$$2\pi n \max_{1\leq k\leq n} Z_k/(\log(n)T_{n+1}) = 2\pi A_n Y_n\,.$$

Or Y_n et A_n converge en probabilité vers 1. Par suite, en écrivant $|Y_n A_n - 1| \leq |A_n - 1|Y_n + |Y_n - 1|$ on déduit que pour $\eta \leq 1$

$$\begin{aligned}
P(|Y_n A_n - 1| \geq 3\eta) &\leq P(|Y_n A_n - 1| \geq \eta(1+\eta) + \eta) \\
&\leq P(|A_n - 1| \geq \eta) + P(|Y_n - 1| \geq \eta)\,. \qquad (3)
\end{aligned}$$

Comme le membre de droite tend vers 0 avec n, on déduit la convergence en probabilité vers 2π ce qui était demandé.