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

M1-MEFM **Année 2010**

Corrigé du D.S. du 10 novembre 2010

Ex 1. *Ball trap (7 points)*

Soit $(A_i)_{i\geq 1}$ une suite d'événements *indépendants*. On note $p_i = P(A_i)$ et on suppose que :

$$a := \sum_{i=1}^{+\infty} p_i < +\infty.$$

Le but de cet exercice est de prouver l'inégalité

$$\forall n, k \in \mathbb{N}^*, \quad P\Big(\sum_{i=1}^{n} \mathbf{1}_{A_i} \geq k\Big) \leq \frac{a^k}{k!}. \qquad (1)$$

La dernière question propose une application de cette inégalité.

1) La variable aléatoire $\sum_{i=1}^n \mathbf{1}_{A_i}$ est majorée [1] par n. Par conséquent pour $k > n$, l'évènement $\{\sum_{i=1}^n \mathbf{1}_{A_i} \geq k\}$ est l'ensemble vide, d'où

$$P\Big(\sum_{i=1}^{n} \mathbf{1}_{A_i} \geq k\Big) = 0 \leq \frac{a^k}{k!}.$$

L'inégalité (1) est ainsi trivialement vérifiée dans le cas $k > n$. On suppose dans la suite $k \leq n$.

2) Si ω est un élément quelconque de $B_{n,k} := \{\omega \in \Omega;\ \sum_{i=1}^n \mathbf{1}_{A_i}(\omega) \geq k\}$, au moins k termes de la somme $\sum_{i=1}^n \mathbf{1}_{A_i}(\omega)$ valent 1 (les seules valeurs possibles pour chacun étant 0 et 1). Donc ω appartient à au moins k des A_i, autrement dit, il existe un ensemble d'indices $G(\omega) \subset \{1, \dots, n\}$, de cardinal k tel que $\omega \in \bigcap_{i\in G(\omega)} A_i$. On en déduit que

$$\omega \in \bigcup_{\substack{F\subset\{1,\dots,n\}\\ \operatorname{card} F = k}} \bigcap_{i\in F} A_i.$$

Cet ensemble ne dépendant pas de ω et ω étant quelconque dans $B_{n,k}$, ceci établit l'inclusion

$$B_{n,k} \subset \bigcup_{\substack{F\subset\{1,\dots,n\}\\ \operatorname{card} F = k}} \bigcap_{i\in F} A_i.$$

1. Maximum atteint lorsque toutes les indicatrices valent 1. Si tous les p_i sont non nuls, l'indépendance des A_i impose que l'intersection des A_i, $1 \leq i \leq n$, soit non vide (pourquoi ?).

3) De cette inclusion on déduit que

$$P(B_{n,k}) \leq P\left(\bigcup_{\substack{F\subset\{1,\dots,n\}\\ \operatorname{card} F=k}} \bigcap_{i\in F} A_i\right),$$

puis par sous-additivité finie [2] de P,

$$P\left(\bigcup_{\substack{F\subset\{1,\dots,n\}\\ \operatorname{card} F=k}} \bigcap_{i\in F} A_i\right) \leq \sum_{\substack{F\subset\{1,\dots,n\}\\ \operatorname{card} F=k}} P\left(\bigcap_{i\in F} A_i\right).$$

Par indépendance des A_i,

$$P\left(\bigcap_{i\in F} A_i\right) = \prod_{i\in F} P(A_i) = \prod_{i\in F} p_i,$$

d'où finalement

$$P(B_{n,k}) \leq \sum_{\substack{F\subset\{1,\dots,n\}\\ \operatorname{card} F=k}} \prod_{i\in F} p_i.$$

4) On note $a_n = \sum_{i=1}^n p_i$. On peut développer a_n^k comme un produit de k parenthèses toutes égales à la somme $p_1+\cdots+p_n$, en formant tous les produits de k facteurs obtenus en choisissant un terme dans chacune de ces parenthèses. Ceci s'écrit :

$$a_n^k = \sum_{(i_1,\dots,i_k)\in\{1,\dots,n\}^k} p_{i_1}\cdots p_{i_k}.$$

Dans cette somme, tous les termes $(p_{i_1}\cdots p_{i_k})$ sont positifs. En laissant tomber tous ceux pour lesquels les indices $i_1,\dots,i_k$ ne sont pas tous distincts, on obtient un minorant de a_n^k. Dans ce minorant, on regroupe par paquets tous les produits $p_{i_1}\cdots p_{i_k}$ ayant les mêmes facteurs à l'ordre près. Chacun des paquets contient $k!$ produits égaux $p_{i_1}\cdots p_{i_k}$ et est associé bijectivement à un sous-ensemble F de cardinal k de $\{1,\dots,n\}$, $F=\{i_1,\dots,i_k\}$. On en déduit :

$$a_n^k \geq k! \sum_{\substack{F\subset\{1,\dots,n\}\\ \operatorname{card} F=k}} \prod_{i\in F} p_i.$$

Nota bene. Cette question a été la plus maltraitée dans les copies. Si la solution ci-dessus ne vous paraît pas claire, prenez votre courage à deux mains, écrivez les 27 termes du développement de $(p_1+p_2+p_3)^3$ et suivez pas à pas le raisonnement ci-dessus sur cet exemple.

2. Attention, contrairement à ce qu'on a pu lire dans de trop nombreuses copies, la réunion indexée par F n'est pas disjointe en général et on ne peut donc pas invoquer l'additivité finie de P ici. Si les p_i sont tous non nuls, on est même certain qu'aucune intersection $(\bigcap_{i\in F} A_i)\cap(\bigcap_{i\in F'} A_i)$ n'est vide, voir la note 1.

5) En combinant les résultats des deux questions précédentes, on obtient

$$P(B_{n,k}) \le \frac{a_n^k}{k!}.$$

Comme les p_i sont positifs, $a_n \le a = \sum_{i=1}^{+\infty} p_i$, d'où pour tout n et tout k entiers (compte-tenu aussi de la question 1),

$$P(B_{n,k}) \le \frac{a^k}{k!}.$$

L'inégalité (1) est démontrée.

6) *Application à un problème de tir.* Dans un stand de tir, une cible mobile traverse le champ visuel d'un tireur (une traversée par épreuve). À chaque épreuve, le tireur tire un coup sur la cible. D'une épreuve à la suivante, la vitesse de la cible augmente de 20%. On suppose que pour un tireur donné, la probabilité de toucher la cible est *inversement proportionnelle* à la vitesse de la cible. Elle vaut ainsi $p \in]0,1[$ pour le premier tir, $\frac{5}{6}p$ pour le second (pourquoi ?), etc. Les tirs sont supposés indépendants, le tireur dispose d'autant de cartouches qu'il le souhaite et le défi qu'il doit relever est celui de toucher au moins 20 fois la cible. En utilisant le résultat démontré ci-dessus, on va majorer sa probabilité de réussir (indépendamment de p).

Notons p_i la probabilité de toucher la cible au i^{e} tir, avec $p_1 = p$ et v_i la vitesse de la cible lors de la i^{e} épreuve. Puisque p_i est inversement proportionnelle à v_i, il existe une constante C telle que pour tout $i \ge 1$, $p_i = C/v_i$. D'autre part la vitesse augmente de 20% d'une épreuve à la suivante, ce qui se traduit par la relation $v_i = 1{,}2\, v_{i-1}$, ou encore $v_i = \frac{6}{5} v_{i-1}$, pour tout $i \ge 2$, d'où

$$\forall i \ge 2, \quad p_i = \frac{C}{v_i} = \frac{5}{6}\frac{C}{v_{i-1}} = \frac{5}{6} p_{i-1}.$$

On voit ainsi que $(p_i)_{i\ge 1}$ est la suite géométrique de premier terme $p_1 = p$ et de raison $\frac{5}{6} \in]-1,1[$. Par conséquent la série de terme général p_i converge et

$$a = \sum_{i=1}^{+\infty} p_i = p \sum_{i=1}^{+\infty} \left(\frac{5}{6}\right)^{i-1} = p \sum_{j=0}^{+\infty} \left(\frac{5}{6}\right)^{j} = \frac{p}{1-\frac{5}{6}} = 6p.$$

Notons A_i l'évènement « le tireur touche la cible lors du i^{e} tir ». Les A_i héritent de l'indépendance des tirs. Les conditions de l'inégalité (1) sont toutes satisfaites. On en déduit que pour tout n,

$$P(B_{n,20}) \le \frac{(6p)^{20}}{20!} \le \frac{6^{20}}{20!} < 0{,}001\,51. \tag{2}$$

Ceci nous fournit un majorant de la probabilité de toucher au moins 20 fois la cible en n tirs consécutifs, quel que soit le nombre n de tirs choisis par le tireur.

On peut améliorer ce résultat en notant que la suite d'évènements $(B_{n,20})_{n\ge 1}$ est croissante pour l'inclusion et donc que $P(B_{n,20})$ converge en croissant quand n tend vers l'infini vers $P(B_{20})$, où l'on note

$$B_{20} := \bigcup_{j\geq 1} B_{j,20} = \left\{ \sum_{i=1}^{+\infty} \mathbf{1}_{A_i} \geq 20 \right\}.$$

On vous laisse le soin de justifier proprement la dernière égalité. En passant à la limite dans (2), on obtient finalement $P(B_{20}) < 0{,}001\,51$. La probabilité de toucher au moins 20 fois la cible est donc majorée par 0,001 51, même si le tireur effectue une infinité de tentatives[3].

Ex 2. *Calibrage de pommes (6 points)*

Une coopérative agricole a un contrat de fourniture de pommes de catégorie A, c'est-à-dire dont le diamètre en mm est dans l'intervalle $[67, 73]$. Le gérant de la coopérative a besoin d'évaluer rapidement[4] la proportion p de pommes *hors catégorie* A dans la récolte qu'il vient d'emmagasiner. Pour cela, il prélève au hasard un échantillon de 400 pommes dont une calibreuse mécanique lui permet d'enregistrer les diamètres. Les données relevées sont présentées dans le tableau suivant. Les diamètres hors catégorie apparaissent en gras.

66	68	71	**66**	68	70	71	69	**65**	70	70	73	68	**75**	69	**65**	69	69	70	73
70	**65**	69	**74**	71	71	72	69	67	71	68	68	70	67	70	**66**	69	**66**	67	71
69	67	**74**	72	71	67	69	67	68	72	68	72	67	**65**	70	72	73	69	68	**66**
66	73	71	**66**	**64**	68	**62**	71	72	67	71	68	68	73	67	71	69	70	73	70
71	67	72	68	72	**62**	72	71	70	72	69	71	67	68	69	72	**66**	69	73	**62**
72	**66**	69	71	**75**	71	70	67	69	69	72	73	**66**	70	70	69	68	70	69	**61**
69	69	69	71	70	69	70	69	67	68	69	69	67	**66**	70	67	71	68	70	68
67	68	72	68	**74**	68	**76**	71	69	70	73	71	68	71	70	73	72	72	**66**	**66**
72	**65**	**66**	69	67	69	**76**	**76**	69	**66**	69	**65**	72	67	**74**	68	67	67	69	68
71	70	72	67	72	67	68	69	70	69	70	69	**74**	72	**74**	68	69	69	69	73
71	70	68	67	73	69	68	72	72	**65**	**66**	67	72	69	69	72	70	70	**66**	71
72	70	70	72	68	73	70	67	69	71	70	67	71	72	67	70	69	**65**	**65**	70
71	73	**66**	67	**65**	69	69	73	71	72	67	68	68	71	68	67	70	72	68	68
69	68	**64**	69	**65**	70	68	70	67	67	70	72	69	72	71	**66**	68	71	68	**65**
73	68	**74**	**76**	**74**	67	71	72	70	**66**	72	67	73	68	**74**	**65**	69	71	70	71
73	**66**	71	68	70	70	**63**	69	70	69	73	69	73	71	69	70	68	**65**	70	68
70	71	71	68	72	71	72	70	**66**	71	68	69	72	**65**	68	67	71	68	72	70
68	71	**74**	72	71	**66**	69	71	71	68	68	**65**	73	70	68	**77**	**61**	73	68	70
68	71	72	69	69	**65**	70	67	72	71	72	69	67	**75**	71	71	69	67	72	67
68	68	70	67	71	67	69	68	70	**65**	70	70	69	71	70	**66**	70	**75**	72	**74**

On dénombre 70 pommes hors catégorie dans cet échantillon de taille 400. À partir de cette observation, nous allons construire deux intervalles de confiance au niveau 95% pour la proportion inconnue p de pommes hors catégorie dans la population totale.

Remarquons d'abord que l'échantillon observé résulte d'un tirage *sans remise* de 400 individus dans une population de grande taille[5] N. En toute rigueur, la loi du nombre S_n

3. Bien sûr, c'est physiquement exclu, puisque la cible dépasserait la vitesse de la lumière ...
4. Avant de lancer l'opération de calibrage et d'emballage.
5. En comptant 6 pommes au kg, cela donnerait environ $N = 600\,000$ pommes pour une récolte de 100 tonnes.

de pommes hors catégorie dans l'échantillon de taille $n = 400$ est donc hypergéométrique de paramètres N, pN et n, mais vu la taille de la population, il est légitime d'approximer cette loi par la loi binomiale de paramètres $n = 400$ et p (inconnue) qui serait celle de S_n si les observations faites résultaient d'un tirage *avec remise*.

Dans toute la suite de l'exercice nous supposerons donc que S_n suit la loi binomiale $\mathrm{Bin}(400, p)$.

Méthode avec variance majorée. En notant

$$S_n^* := \frac{S_n - \mathbf{E}S_n}{\sqrt{\operatorname{Var} S_n}} = \frac{S_n - np}{\sqrt{np(1-p)}} = \sqrt{\frac{n}{p(1-p)}}\left(\frac{S_n}{n} - p\right),$$

la somme centrée réduite, le théorème de de Moivre-Laplace nous dit que S_n^* converge en loi quand n tend vers l'infini vers Z de loi normale $\mathfrak{N}(0,1)$. La fonction de répartition Φ de Z étant *continue* sur $\mathbb{R}$, on en déduit que

$$\forall t > 0, \quad P(|S_n^*| \leq t) \xrightarrow[n\to+\infty]{} P(|Z| \leq t) = \Phi(t) - \Phi(-t) = 2\Phi(t) - 1,$$

en utilisant pour la dernière égalité la symétrie de la loi de Z. On peut réécrire cette convergence sous la forme :

$$\forall n \geq 1, \forall t > 0, \quad P(-t \leq S_n^* \leq t) = 2\Phi(t) - 1 + \varepsilon_n(t), \quad \varepsilon_n(t) \xrightarrow[n\to+\infty]{} 0,$$

où $\varepsilon_n(t)$ est l'erreur d'approximation gaussienne.

En résolvant par rapport à p l'encadrement $-t \leq S_n^* \leq t$, on voit que

$$-t \leq S_n^* \leq t \quad \Longleftrightarrow \quad \frac{S_n}{n} - t\sqrt{\frac{p(1-p)}{n}} \leq p \leq \frac{S_n}{n} + t\sqrt{\frac{p(1-p)}{n}}.$$

Cet encadrement n'est pas satisfaisant pour construire un intervalle de confiance pour p car les bornes dépendent de p via la quantité inconnue $p(1-p)$ qui est la variance d'une v.a. de Bernoulli de paramètre p. La méthode avec variance majorée consiste à se débarasser de cet inconvénient en notant que la fonction $g : p \mapsto p(1-p)$ atteint son maximum sur $[0,1]$ au point $p_0 = \frac{1}{2}$ et est donc majorée par $g(p_0) = \frac{1}{2}\left(1 - \frac{1}{2}\right) = \frac{1}{4}$. Ainsi pour tout $p \in [0,1]$, $\sqrt{p(1-p)} \leq \sqrt{\frac{1}{4}} = \frac{1}{2}$, d'où :

$$-t \leq S_n^* \leq t \quad \Longrightarrow \quad \frac{S_n}{n} - \frac{t}{2\sqrt{n}} \leq p \leq \frac{S_n}{n} + \frac{t}{2\sqrt{n}}.$$

Cette implication se traduit par l'inclusion d'évènements $A_{n,t} \subset B_{n,t}$, en notant

$$A_{n,t} = \{-t \leq S_n^* \leq t\}, \quad B_{n,t} = \left\{\frac{S_n}{n} - \frac{t}{2\sqrt{n}} \leq p \leq \frac{S_n}{n} + \frac{t}{2\sqrt{n}}\right\}.$$

On en déduit que

$$P(B_{n,t}) \geq P(A_{n,t}) = 2\Phi(t) - 1 + \varepsilon_n(t).$$

On cherche t tel que $2\Phi(t)-1 = 0{,}95$, c'est-à-dire $\Phi(t) = 0{,}975$, d'où $t = 1{,}96$ (par lecture inverse de la table des valeurs de Φ). Considérant alors que $n = 400$ est suffisamment grand pour que l'on puisse négliger l'erreur d'approximation gaussienne $\varepsilon_n(t)$, on obtient finalement $P(B_{n,t}) \geq 0{,}95$ pour $n = 400$ et $t = 1{,}96$.

En réalité, ce que l'on a observé, c'est *une réalisation particulière* $S_n(\omega) = 70$ de la variable aléatoire S_n et comme on ne connaît pas p, on ne peut pas dire avec certitude si le ω sous-jacent appartient ou non à $B_{n,t}$. On « parie » donc sur la réalisation de $B_{n,t}$ et d'après l'étude précédente, la probabilité de gagner ce pari est (approximativement) au moins 95%. L'intervalle I d'encadrement de p correspondant à la réalisation de $B_{n,t}$ avec $S_n(\omega) = 70$ est un intervalle de confiance pour p au niveau 95%.

Numériquement on obtient :

$$I = \left[\frac{70}{400} - \frac{1{,}96}{2\sqrt{400}}; \frac{70}{400} + \frac{1{,}96}{2\sqrt{400}}\right] = [0{,}126; 0{,}224].$$

Méthode avec variance estimée. Au lieu de majorer $p(1-p)$, on l'estime par $V_n = \frac{S_n}{n}(1 - \frac{S_n}{n})$. V_n converge presque sûrement vers $p(1-p)$ grâce à la loi forte des grands nombres et à la conservation de la convergence presque sûre par l'application continue $g : [0,1] \to \mathbb{R}$, $x \mapsto x(1-x)$. En notant

$$C_{n,t} = \left\{ \frac{S_n}{n} - \frac{t\sqrt{V_n}}{\sqrt{n}} \leq p \leq \frac{S_n}{n} + \frac{t\sqrt{V_n}}{\sqrt{n}} \right\},$$

on obtient

$$P(C_{n,t}) = 2\Phi(t) - 1 + \varepsilon'_n(t)$$

puis avec le choix $t = 1{,}96$ et en négligeant l'erreur d'approximation gaussienne $\varepsilon'_n(t)$, $P(C_{n,t}) \simeq 0{,}95$. L'intervalle de confiance correspondant est

$$J = [0{,}137; 0{,}213].$$

La justification de ce procédé repose sur le théorème suivant.

Théorème 1 (TLC avec autonormalisation)**.** *Soient $X_1, \dots, X_n, \dots$ des variables aléatoires indépendantes et de même loi telle que $\mathbf{E}X_1^2 < +\infty$ et $\sigma^2 := \operatorname{Var} X_1 > 0$. On note $S_n := X_1 + \dots + X_n$. On suppose de plus que $(V_n)_{n\geq 1}$ est une suite de variables aléatoires positives qui converge en probabilité vers σ^2. Alors*

$$T_n := \sqrt{\frac{n}{V_n}}\left(\frac{S_n}{n} - \mathbf{E}X_1\right) \xrightarrow[n\to+\infty]{\text{loi}} Z,$$

où Z suit la loi gaussienne standard $\mathfrak{N}(0,1)$.

Ex 3. *Transformée de Laplace (7 points)*

La transformée de Laplace d'une variable aléatoire positive X est la fonction L_X définie sur $\mathbb{R}_+$ par :

$$\forall t \in \mathbb{R}_+, \quad L_X(t) = \mathbf{E}\exp(-tX).$$

Clairement L_X ne dépend que de la loi de X. On parlera donc aussi bien de la transformée de Laplace d'une variable que de sa loi.

Remarque simplificatrice. On convient une fois pour toutes que la *positivité* de la variable aléatoire X, signifie que $X(\omega) \geq 0$ *pour tout* $\omega \in \Omega$. Comme X n'intervient que via des espérances de la forme $\mathbf{E}h(X)$, les résultats obtenus dans cet exercice s'étendent facilement au cas d'une variable aléatoire X p.s. positive.

1) Pour tout $t \in \mathbb{R}$, $\exp(-tX)$ est une variable aléatoire positive, donc son espérance existe dans $\overline{\mathbb{R}}_+$. De plus comme X est une v.a. positive, $0 \leq \exp(-tX) \leq 1$, pour tout $t \in \mathbb{R}_+$. Dans ce cas, la variable $\exp(-tX)$ est bornée, donc intégrable et $L_X(t)$ est un nombre réel positif. La fonction L_X définit donc bien une application $\mathbb{R}_+ \to \mathbb{R}_+$.
Remarque. Pour certaines variables aléatoires positives, l'ensemble de définition de la fonction L_X peut être plus grand que $[0, +\infty[$. Par exemple, si X suit une loi exponentielle de paramètre a, L_X est définie sur $]-a, +\infty[$ (vérifiez).

Si X suit la loi de Bernoulli de paramètre p, on a pour toute fonction h définie au moins sur $\{0, 1\}$,

$$\mathbf{E}h(X) = h(0)P(X = 0) + h(1)P(X = 1) = (1 - p)h(0) + ph(1).$$

En appliquant ceci avec $h(x) = \exp(-tx)$ (avec t fixé), on obtient

$$L_X(t) = \mathbf{E}\exp(-tX) = 1 - p + p\exp(-t).$$

Cette formule est valide pour tout t réel (notez que si X suit une loi de Bernoulli, $\exp(-tX)$ est une variable aléatoire positive bornée par $\exp(|t|)$). Ceci nous fournit un exemple de transformée de Laplace définie sur $\mathbb{R}$ tout entier. De plus il est clair que L_X est décroissante et convexe et a pour limite 1 quand t tend vers 0 et $1 - p = P(X = 0)$ quand t tend vers $+\infty$.

2) Revenant au cas général d'une v.a. positive X quelconque, la décroissance de L_X se vérifie comme suit. On note d'abord que pour tout réel $x \geq 0$, la fonction $t \mapsto \exp(-xt)$ est décroissante sur $\mathbb{R}_+$. Par conséquent si $0 \leq s < t$, l'inégalité $\exp(-sX(\omega)) \geq \exp(-tX(\omega))$ est vérifiée pour tout $\omega \in \Omega$. On peut donc la réécrire comme l'inégalité entre variables aléatoires $\exp(-sX) \geq \exp(-tX)$. Par croissance de l'espérance, on en déduit $\mathbf{E}\exp(-sX) \geq \mathbf{E}\exp(-tX)$, autrement dit $L_X(s) \geq L_X(t)$. Ceci étant valide pour tous réels positifs s et t tels que $s < t$, la décroissance de la fonction L_X sur $\mathbb{R}_+$ est établie.

3) Vérifions la convexité de L_X sur $\mathbb{R}_+$. On rappelle que f est convexe sur $\mathbb{R}_+$ si pour tout $u \in [0, 1]$, et tous $s, t \in \mathbb{R}_+$, $f(us + (1 - u)t) \leq uf(s) + (1 - u)f(t)$. Il est bien connu que la fonction exponentielle est convexe sur $\mathbb{R}$. On en déduit immédiatement que la fonction $t \mapsto \exp(-xt)$ est convexe sur $\mathbb{R}$ pour tout réel x (le fait que la corde soit toujours au dessus de la courbe est une propriété géométrique qui n'est pas affectée par le changement de paramétrage $t \mapsto -xt$ en abcsisse). On en déduit que pour tout $u \in [0, 1]$, et tous $s, t \in \mathbb{R}_+$,

$$\forall \omega \in \Omega, \quad \exp\Big(-(us + (1 - u)t)X(\omega)\Big) \leq u\exp\Big(-sX(\omega)\Big) + (1 - u)\exp\Big(-tX(\omega)\Big).$$

On peut réécrire ceci comme inégalité entre variables aléatoires

$$\exp\Big(-\Big(us+(1-u)t\Big)X\Big) \leq u\exp(-sX)+(1-u)\exp(-tX),$$

prendre l'espérance des deux membres et en déduire par croissance et linéarité de l' espérance :

$$\begin{aligned}\mathbf{E}\exp\Big(-\Big(us+(1-u)t\Big)X\Big) &\leq \mathbf{E}\Big(u\exp(-sX)+(1-u)\exp(-tX)\Big)\\ &= u\mathbf{E}\exp(-sX)+(1-u)\mathbf{E}\exp(-tX)\end{aligned}$$

Comme $s,t\in\mathbb{R}_+$ et $u\in[0,1]$ étaient quelconques, la convexité de L_X est vérifiée.

Note du correcteur. Dans certaines copies, j'ai pu voir une tentative de prouver la convexité de L_X en vérifiant la positivité de L_X''. Si on choisit cette voie (inutilement compliquée), il faut justifier la dérivabilité à l'ordre 2 de L_X (ceux qui connaissent l'intégrale de Lebesgue y verront un exercice sur le théorème de dérivation sous le signe somme, qui est une conséquence du théorème de convergence dominée), ce qui revient en pratique à vérifier l'intégrabilité de $X^2\exp(-tX)$. Vous pouvez utiliser un théorème hors programme, mais dans ce cas il importe de le citer avec précision, ce qui n'était fait dans aucune copie.

4) La convexité de L_X sur $\mathbb{R}_+$ implique sa continuité sur $]0,+\infty[$. Pour montrer la continuité de L_X au point 0, il suffit, compte-tenu de la monotonie de L_X, de montrer que $L_X(1/n)$ tend vers $L_X(0)=1$ quand n tend vers $+\infty$. Justifions cette convergence à l'aide du théorème de B. Levi.

L'image d'une suite décroissante de réels par une fonction décroissante est une suite croissante de réels. En appliquant ceci avec la suite $(1/n)_{n\geq1}$ et la fonction $t\mapsto\exp(-xt)$ (avec $x\geq0$ quelconque mais fixé), on voit que la suite de réels $(\exp(-x/n))_{n\geq1}$ est croissante pour tout $x\in\mathbb{R}_+$. En prenant $x=X(\omega)$ avec ω quelconque, on en déduit que la suite de variables aléatoires $(Y_n)_{n\geq1}=(\exp(-X/n))_{n\geq1}$ est *croissante*. Elle converge vers la v.a. constante $\exp(0)=1$ par continuité de l'exponentielle. Comme ces variables aléatoires sont aussi positives, le théorème de B. Levi nous dit que $\mathbf{E}Y_n$ converge en croissant vers $\mathbf{E}(\lim_{n\to+\infty}Y_n)=\mathbf{E}1=1$. Ainsi $L_X(1/n)$ converge vers $1=L_X(0)$ quand n tend vers l'infini. Et par monotonie de L_X, on en déduit que $L_X(t)$ converge vers $L_X(0)$ quand t tend vers 0. Ainsi L_X est bien continue au point 0.
Remarque. Comme $0\leq Y_n\leq1$, on pouvait aussi appliquer le théorème de convergence dominée au lieu de celui de B. Levi.

5) Par une argumentation similaire à celle de la question précédente, on obtient la décroissance de la suite de variables aléatoires $Y_n=\exp(-nX)$. La limite de $Y_n(\omega)$ lorsque n tend vers l'infini s'obtient en distinguant deux cas.
Cas 1. Si $X(\omega)>0$, $-nX(\omega)$ tend vers $-\infty$ et son exponentielle $Y_n(\omega)$ tend vers 0.
Cas 2. Si $X(\omega)=0$, $Y_n(\omega)=\exp(0)=1$ pour tout n et la limite est trivialement 1.

On voit ainsi que la suite de variables aléatoires $(Y_n)_{n\geq1}$ converge sur tout Ω vers $\mathbf{1}_{\{X=0\}}$. On ne peut pas utiliser directement le théorème de B. Levi pour intervertir limite et espérance, puisque ce théorème concerne les suites *croissantes* de variables aléatoires positives. Néanmoins puisque $0 \leq Y_n \leq 1$, en posant $Z_n = 1 - Y_n$, on retrouve une suite croissante de v.a. positives et en lui appliquant le théorème de B. Levi,

$$1 - \mathbf{E}Y_n = \mathbf{E}Z_n \uparrow \mathbf{E}(1 - \mathbf{1}_{\{X=0\}}) = 1 - \mathbf{E}\mathbf{1}_{\{X=0\}},$$

on obtient la convergence de $\mathbf{E}Y_n$ vers $\mathbf{E}\mathbf{1}_{\{X=0\}} = P(X = 0)$. On a donc montré la convergence de $L_X(n)$ vers $P(X = 0)$ quand n tend vers l'infini. En raison de la monotonie de L_X on en déduit que $L_X(t)$ tend vers $P(X = 0)$ quand t tend vers $+\infty$.

Là encore, on pouvait appliquer alternativement le théorème de convergence dominée pour obtenir la convergence de $\mathbf{E}Y_n$.

Commentaire sur l'interversion limite espérance pour une suite décroissante. Dans cette question, comme dans la précédente, l'utilisation de la monotonie de la suite (Y_n) était un chemin moins rapide que le théorème de convergence dominée. Ce détour est l'occasion d'attirer votre attention sur le caractère non symétrique du théorème de B. Levi dont l'énoncé devient faux si on se contente de remplacer la suite croissante de v.a. positives par une suite décroissante de v.a. positives. Voici un contre-exemple. On prend X suivant la loi de Cauchy et $Y_n = |X|\mathbf{1}_{\{|X|>n\}}$. Alors (Y_n) est décroissante et a pour limite 0, mais $\mathbf{E}Y_n = +\infty$ pour tout n et ne peut donc converger vers $\mathbf{E}0 = 0$.

6) Si X et Y sont des variables aléatoires indépendantes, leurs images respectives par des applications boréliennes g, h sont des variables aléatoires $g(X)$ et $h(Y)$ indépendantes. En choisissant pour g et h l'application continue donc borélienne $x \mapsto \exp(-tx)$, avec t fixé, on obtient l'indépendance des variables aléatoires positives $\exp(-tX)$ et $\exp(-tY)$. Ceci légitime le calcul suivant, :

$$\begin{aligned} L_{X+Y}(t) = \mathbf{E}\exp\Big(-t(X+Y)\Big) &= \mathbf{E}\Big(\exp(-tX)\exp(-tY)\Big) \\ &= \mathbf{E}\exp(-tX)\mathbf{E}\exp(-tY) = L_X(t)L_Y(t). \end{aligned}$$

Comme ce calcul est valide pour tout $t \in \mathbb{R}_+$, il établit l'égalité fonctionnelle $L_{X+Y} = L_X L_Y$ vraie pour toute paire de v.a. positives et indépendantes X et Y.

7) Pour estimer $L_X(t)$, à partir de l'observation d'un échantillon $X_1, \dots, X_n$ de variables aléatoires indépendantes de même loi (inconnue) que X, on peut prendre

$$M_n(t) := \frac{1}{n}\sum_{i=1}^{n} \exp(-tX_i).$$

Cet estimateur est *sans biais* puisque les X_i ayant même loi, il en va de même pour leur image par la même fonction continue (donc borélienne) $x \mapsto \exp(-tx)$. En particulier chaque $\exp(-tX_i)$ a même loi que $\exp(-tX)$ et donc aussi même espérance puisque $\exp(-tX)$ est intégrable. Par conséquent, par linéarité de l'espérance,

$$\mathbf{E}M_n(t) = \frac{1}{n}\sum_{i=1}^{n} \mathbf{E}\exp(-tX_i) = \frac{1}{n}(n\mathbf{E}\exp(-tX)) = \mathbf{E}\exp(-tX) = L_X(t).$$

On vient de voir que les $\exp(-tX_i)$ ont même loi et sont intégrables. Comme elles héritent de l'indépendance des X_i (images des X_i par une application borélienne), on peut leur appliquer la loi forte des grands nombres, qui s'écrit ici :

$$M_n(t) = \frac{1}{n}\sum_{i=1}^{n} \exp(-tX_i) \xrightarrow[n\to+\infty]{\text{p.s.}} \mathbf{E}\exp(-tX) = L_X(t).$$

$M_n(t)$ est donc un estimateur fortement consistant de $L_X(t)$.

Un dernier commentaire [6]. La transformée de Laplace a des propriétés analogues à celles de la fonction de répartition (en plus simple mais moins universel). En particulier $L_X = L_Y$ si et seulement si X et Y ont même loi et Y_n converge en loi vers Y si et seulement si $L_{Y_n}(t)$ converge vers $L_Y(t)$ pour tout t dans $\mathbb{R}_+$. L'estimateur M_n ci-dessus est la transformée de Laplace empirique. La transformée de Laplace est très utilisée en fiabilité et dans les mathématiques du risque (assurances, ...).

6. C'est promis.