Corrigé de l'examen de Probabilités 1ère session - 7 juin 2007

Exercice 1 (Question de cours, 3 points). Soit $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires de $(\Omega,\mathcal{A},\mathbb{P})$ dans $(\mathbb{R},\mathcal{B}(\mathbb{R}))$ et $X:(\Omega,\mathcal{A},\mathbb{P})\to(\mathbb{R},\mathcal{B}(\mathbb{R}))$ une autre variable aléatoire.
a) Donner la définition de la convergence presque sûre de (X_n) vers X. Donner un exemple d'une telle suite $(X_n)_{n\in\mathbb{N}}$ convergente
b) Donner la définition de la convergence en probabilité de (X_n) vers X. Donner un exemple de suite $(X_n)_{n\in\mathbb{N}}$ qui converge en probabilité mais pas presque surement vers X.
c) Donner la définition de la convergence en loi de (X_n) vers X. Donner un exemple de suite $(X_n)_{n\in\mathbb{N}}$ qui converge en loi mais pas en probabilité vers X.

Exercice 2 (Un exercice fait en TD, 4 points). Soit (X_n) une suite de variables aléatoires de $(\Omega,\mathcal{A},\mathbb{P})$ dans $(\mathbb{R},\mathcal{B}(\mathbb{R}))$.
a) Montrer que si $X_n\to X$ en probabilité, alors $E(\min(|X_n-X|,1)\to 0$ quand $n\to\infty$.
b) Montrer que si $E(\min(|X_n-X|,1)\to 0$ quand $n\to\infty$, alors $X_n\to X$ en probabilité.
c) À l'aide de **a)**, montrer que si $X_n\to X$ en probabilité, alors il existe une sous-suite (X_{n_k}) qui tend vers X presque sûrement.

Corrigé : a) On sait que pour tout $\varepsilon>0$ fixé, on a $P(|X_n-X|\geq\varepsilon)\to 0$ quand $n\to\infty$. On majore

$$\begin{aligned}E(\min(|X_n-X|,1)) &= \int_{|X_n-X|\geq\varepsilon}\min(|X_n-X|,1)\,d\mathbb{P}+\int_{|X_n-X|<\varepsilon}\min(|X_n-X|,1)\,d\mathbb{P}\\ &\leq 1.P(|X_n-X|\geq\varepsilon)+\varepsilon P(|X_n-X|<\varepsilon\leq P(|X_n-X|\geq\varepsilon)+\varepsilon\,.\end{aligned}$$

Quand $n\to\infty$ on en déduit que $0\leq\limsup_{n\to\infty}E(\min(|X_n-X|,1))\leq\varepsilon$. Ceci étant vrai pour tout $\varepsilon>0$, on en déduit que $E(\min(|X_n-X|,1)\to 0$ quand $n\to\infty$.
b) Réciproquement, notons d'abord que pour $\varepsilon=1$, on a $E(\min(|X_n-X|,1))\geq P(|X_n-X|\geq 1)$. Si $E(\min(|X_n-X|,1))\to 0$, alors $P(|X_n-X|\geq 1)\to 0$ quand $n\to\infty$. On cherche à montrer que pour tout $\varepsilon>0$, $P(|X_n-X|\geq\varepsilon)\to 0$ quand $n\to\infty$. On sait le montrer pour $\varepsilon=1$. Il suffit donc de montrer que pour tout $0<\varepsilon<1$, $P(\varepsilon\leq|X_n-X|<1)\to 0$ quand $n\to\infty$. Or $P(\varepsilon\leq|X_n-X|<1)=P(\varepsilon\leq\min(|X_n-X|,1)<1)\leq\frac{1}{\varepsilon}E(\min(|X_n-X|,1))$ d'après l'inégalité de Markov. Donc $P(\varepsilon\leq|X_n-X|<1)\to 0$ quand $n\to\infty$.
c) Supposons que $X_n\to X$ en probabilité. D'après a), $E(\min(|X_n-X|,1)\to 0$ quand $n\to\infty$. On peut donc construire une suite strictement croissante $(n_k)_{k\geq 1}$ d'indices tels que pour tout $k\geq 1$, $E(\min(|X_{n_k}-X|,1)\leq\frac{1}{2^k}$. Ceci implique bien sûr que la série $\sum_{k\geq 1}E(\min(|X_{n_k}-X|,1))$ converge. Or $\sum_{k=1}^{+\infty}E(\min(|X_{n_k}-X|,1))=\lim_{K\to\infty}E(\sum_{k=1}^{K}\min(|X_{n_k}-X|,1))$. Comme $\min(|X_{n_k}-X|,1)\geq 0$, la suite de fonctions $(\sum_{k=1}^{K}\min(|X_{n_k}-X|,1))_{K\geq 1}$ est une suite croissante de fonctions positives. Le théorème de convergence monotone donne $\lim_{K\to\infty}E(\sum_{k=1}^{K}\min(|X_{n_k}-X|,1))=E(\lim_{K\to\infty}\sum_{k=1}^{K}\min(|X_{n_k}-X|,1))=E(\sum_{k=1}^{+\infty}\min(|X_{n_k}-X|,1))$. On déduit de tout ce qui précède que la quantité $E(\sum_{k=1}^{+\infty}\min(|X_{n_k}-X|,1))$ est finie. Cette intégrale étant finie, ceci implique que la série $\sum_{k=1}^{+\infty}\min(|X_{n_k}-X|,1))$ est finie p.s. Donc son terme général $\min(|X_{n_k}-X|,1))$ tend vers 0 p.s. quand $k\to\infty$. En particulier, il existe un rank $k_0\geq 1$ à partir duquel on a $\min(|X_{n_k}-X|,1))<1$, et donc $\min(|X_{n_k}-X|,1))=|X_{n_k}-X|$. Donc $|X_{n_k}-X|$ tend vers 0 p.s. quand $k\to\infty$.
d) Si $X_n\to X$ en proba, alors $X_{n_k}\to X$ en proba aussi. Donc il existe une sous-suite $(X_{n_{k_j}})$ qui converge ps vers X.
Supposons maintenant que X_n ne converge pas vers X en proba. Alors il existe $\varepsilon>0$, $\alpha>0$ et une suite strictement croissante $(n_k)_{k\in\mathbb{N}}$ tq pour tout $k\in\mathbb{N}$ $P(|X_{n_k}-X|\geq\varepsilon)\geq\alpha>0$. Donc aucune sous-suite de X_{n_k} ne peut converger vers X p.s.

Exercice 3 (5 points). On considère une population de n individus numérotés $1,\dots,n$ ($n\geq 2$ est très grand). À chaque individu i est associé son revenu mensuel $r_i\geq 0$. On note $\bar{r}=\frac{1}{n}\sum_{i=1}^{n}r_i$ le revenu moyen de la population.
a) On choisit un individu ω au hasard suivant une loi uniforme (notée $\mathbb{P}$) sur $\Omega=\{1,\dots,n\}$. Soit $1\leq i\leq n$ un individu fixé. Quelle est la probabilité de choisir l'individu i ?
b) On note $Y:(\Omega,\mathcal{P}(\Omega),\mathbb{P})\to(\mathbb{R}_+,\mathcal{B}(\mathbb{R}_+))$ la variable aléatoire qui à chaque individu i associe son revenu r_i. Justifier le fait que Y est une variable aléatoire.
c) Calculer l'espérance $E(Y)$ et la variance $\mathrm{Var}(Y)$.
d) On sonde maintenant la population en choisissant k fois de suite ($k\geq 1$) un individu au hasard sur Ω. On suppose que les k choix sont indépendants et faits suivant la loi uniforme notée $\mathbb{P}$ sur Ω. (Noter qu'un individu peut donc être choisi plusieurs fois.) On note $Y_1,Y_2,\dots Y_k$ les revenus des personnes choisies. Calculer l'espérance et la variance de la variable aléatoire $\bar{Y}_k=\frac{1}{k}\sum_{j=1}^{k}Y_j$ en fonction des $(r_i)_{1\leq i\leq n}$ et de $\bar{r}$, ou encore en fonction de $E(Y)$ et $\mathrm{Var}(Y)$.
e) On fait maintenant tendre k vers $+\infty$. Étudier le comportement asymptotique de $E(\bar{Y}_k)$ et $\mathrm{Var}(\bar{Y})_k$ quand $k\to+\infty$.
f) En déduire que $\bar{Y}_k$ converge en probabilité vers une limite que l'on précisera.
g) Retrouver et améliorer le résultat précédent à l'aide d'un résultat du cours que l'on citera précisément.

CORRIGÉ : a) Par définition de la loi uniforme, $\mathbb{P}(i)=\frac{1}{\#\Omega}=\frac{1}{n}$.
b) Soit $B\in\mathcal{B}(\mathbb{R}_+)$ un borélien. $Y^{-1}(B)$ est une partie de Ω, c'est donc un élément de la tribu $\mathcal{P}(\Omega)$ de toutes les parties de Ω. Donc Y est mesurable. C'est donc une variable aléatoire.
c) $E(Y)=\sum_{i=1}^{n}Y(i)\mathbb{P}(\omega=i)=\frac{1}{n}\sum_{i=1}^{n}r_i=\bar{r}$, $E((Y-E(Y))^2)=E((Y-\bar{r})^2)=\sum_{i=1}^{n}(Y(i)-\bar{r})^2\mathbb{P}(\omega=i)=\frac{1}{n}\sum_{i=1}^{n}(r_i-\bar{r})^2$.
d) L'espérance étant linéaire, on a $E(\bar{Y}_k)=\frac{1}{k}\sum_{j=1}^{k}E(Y_j)=\bar{r}$. De plus, les Y_j sont indépendants donc $\mathrm{Var}(\bar{Y}_k)=\frac{1}{k^2}\sum_{j=1}^{k}\mathrm{Var}(Y_j)=\frac{1}{k}\mathrm{Var}(Y)=\frac{1}{k}\frac{1}{n}\sum_{i=1}^{n}(r_i-\bar{r})^2$.
e) Lorsque $k\to\infty$, on a $E(\bar{Y}_k)=E(Y)=\bar{r}$ constante et $\mathrm{Var}(\bar{Y}_k)=\dfrac{\mathrm{Var(Y)}}{k}\to 0$ quand $k\to\infty$.
f) L'inégalité de Bienaymé Tchebytchev donne $P(|\bar{Y}_k-\bar{r}|\geq\varepsilon)\leq\frac{\mathrm{Var}(\bar{Y}_k)}{\varepsilon^2}\to 0$ quand $k\to\infty$. Ceci étant vrai pour tout $\varepsilon>0$, cela signifie que $\bar{Y}_k\to\bar{r}$ en probabilité.
g) Le résultat précédent est donné par la loi faible des grands nombres. Puisque les (Y_i) sont indépendants et de même loi, et $E(|Y_i|)=E(Y_i)<\infty$, la loi forte des grands nombres (voir cours) s'applique et donne la convergence presque sûre des $\bar{Y}_k$ vers $\bar{r}$ quand $k\to\infty$.

Exercice 4 (5 points). Soit $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires indépendantes dans leur ensemble, de même loi exponentielle de paramètre 1. Pour tout $n \geq 1$, on pose $Y_n = X_n + X_{n+1}$. On rappelle que la somme de deux variables aléatoires indépendantes à densités respectives f_1 et f_2 a pour densité le produit de convolution $f_1 \star f_2$ de f_1 et f_2 défini par $f_1 \star f_2(z) = \int_{\mathbb{R}} f_1(z-y)f_2(y)dy = \int_{\mathbb{R}} f_1(y)f_2(z-y)dy$. **a)** Montrer que pour tout $n \geq 1$, Y_n suit la loi $\Gamma(2,1)$ de densité $f(z) = \mathbf{1}_{\mathbb{R}_+}(z)ze^{-z}$.
b) Montrer que pour tout $\alpha > 0$, et tout $n \geq 1$, on a $P(Y_n \geq \alpha) = (1+\alpha)e^{-\alpha}$. **c)** Soit $\beta > 1$ et A_n l'événement $A_n = \{Y_n \geq \beta \ln n\}$. Montrer que $P(\limsup_{n\to\infty}\{Y_n \geq \beta \ln n\}) = 0$. **d)** Pour tout $n \geq 1$, on pose $Z_n = Y_{2n}$. Montrer que $(Z_n)_{n\geq 1}$ est une suite de variables aléatoires indépendantes et de même loi. On précisera leur loi.
e) Montrer que $P(\limsup_{n\to\infty}\{Z_n \geq ln(2n)\} = 1$. **f)** En déduire que presque sûrement, on a $\limsup_{n\to\infty} \frac{Y_n}{\ln n} = 1$. **g)** Pourquoi a-t-on utilisé la suite $(Z_n)_{n\geq 1}$ plutôt que la suite $(Y_n)_{n\geq 1}$ dans les questions d) et e) ?

CORRIGÉ : a) Pour tout $n \geq 1$, X_n et X_{n+1} sont indépendantes et de même loi, à densité $x \mapsto \mathbf{1}_{\mathbb{R}_+}(x)e^{-x}$. On calcule donc $f(z) = \int_{\mathbb{R}} \mathbf{1}_{\mathbb{R}_+}(z-y)\mathbf{1}_{\mathbb{R}_+}(y)e^{-z+y}e^{-y}dy = \mathbf{1}_{\mathbb{R}_+}(z)\int_0^z e^{-z}dy = \mathbf{1}_{\mathbb{R}_+}(z)ze^{-z}$.
b) Soit $\alpha > 0$. On a $P(Y_n \geq \alpha) = \int_\alpha^\infty ze^{-z}dz = \cdots = (1+\alpha)e^{-\alpha}$ (intégration par parties).
c) D'après la question ci-dessus, on a $P(Y_n \geq \beta \ln n) = \frac{(1+\beta \ln n)}{n^\beta}$. Comme $\beta > 1$, on en déduit que la série $\sum_{n\geq 1} P(\{Y_n \geq \beta \ln n\})$ converge. Le théorème de Borel Cantelli (sens facile) donne alors $P(\limsup_{n\to\infty}\{Y_n \geq \beta \ln n\}) = 0$.
d) Remarquons que $Z_n = Y_{2n} = X_{2n} + X_{2n+1}$ et $Z_{n+1} = X_{2n+2} + X_{2n+3}$. Les (X_n) étant indépendants dans leur ensemble, les variables aléatoires Z_n le sont aussi. Et elles sont de même loi puisque les Y_n sont de même loi. Leur loi commune a pour densité f (voir question a).
e) On calcule $P(Z_n \geq \ln 2n) = \frac{1+\ln 2n}{2n}$. C'est le terme général d'une série divergente. Comme les (Z_n) sont indépendantes, on peut utiliser le sens “difficile” de Borel-Cantelli, qui donne $P(\limsup_{n\to\infty}\{Z_n \geq \ln 2n\}) = 1$. Ceci signifie qu'il existe un ensemble $\Omega' \subset \Omega$ de probabilité 1 tq pour tout $\omega \in \Omega'$, il existe une infinité d'entiers, notés $(n_k)_{k\geq 1}$ tq $Z_{n_k}(\omega) \geq \ln(2n_k)$. Ceci se réécrit $Y_{2n_k}(\omega) \geq \ln(2n_k)$. On en déduit que pour tout $\omega \in \Omega'$, on a $\limsup_{n\to\infty} \frac{Y_n}{\ln n} \geq 1$.
Pour montrer l'inégalité inverse, fixons $\beta > 1$. D'après c), pour presque tout ω, il n'existe qu'un nombre fini d'indices n tels que $\frac{Y_n}{\ln n} \geq \beta$. Cela signifie donc que pour presque tout ω, il existe un entier N tel que $\frac{Y_n}{\ln n} \leq \beta$ pour tout $n \geq N$. Ainsi, pour presque tout ω, $\limsup_{n\to\infty} \frac{Y_n(\omega)}{\ln n} \leq \beta$. Cela étant vrai pour tout $\beta > 1$, on déduit que presque surement, $\limsup_{n\to\infty} \frac{Y_n(\omega)}{\ln n} \leq 1$ et donc que $\limsup_{n\to\infty} \frac{Y_n}{\ln n} = 1$ presque surement.
f) On aurait pu vouloir utiliser directement la suite (Y_n), puisque c'est elle qu'on étudie finalement. Mais $Y_n = X_n + X_{n+1}$ et $Y_{n+1} = X_{n+1} + X_{n+2}$. Les Y_n ne sont donc pas indépendants (ni deux à deux ni dans leur ensemble). On n'aurait donc pas pu utiliser la partie “difficile” de Borel cantelli avec la suite Y_n à la question e).

Exercice 5 (6 points). Première partie Soit $X : (\Omega, \mathcal{A}, \mathbb{P}) \to (\mathbb{R}_+, \mathcal{B}(\mathbb{R}_+))$ une variable aléatoire de loi exponentielle de paramètre $\lambda > 0$.
a) Donner sa densité ainsi que l'allure de son graphe.
b) Donner sa fonction de répartition ainsi que l'allure de son graphe.
c) Donner sa fonction caractéristique.
d) Mêmes questions (densité et fonction caractéristique) si Z est une variable aléatoire de loi normale $\mathcal{N}(0,1)$.
e) Mêmes questions (densité et fonction caractéristique) si Z est une variable aléatoire de loi normale $\mathcal{N}(m,\sigma^2)$ d'espérance m et de variance σ^2.

Deuxième partie Soient $X_1, Y_1, X_2, Y_2, \ldots, X_n, Y_n, \ldots$ une suite de variables aléatoires réelles positives indépendantes suivant toutes la même loi de probabilité exponentielle de paramètre $\lambda > 0$.
f) Calculer la fonction caractéristique φ de la variable aléatoire $X_1 - Y_1$.
g) À l'aide d'un développement limité de $\varphi(t)$, en déduire les valeurs de $E(X_1 - Y_1)$ et $E((X_1 - Y_1)^2)$.
h) Retrouver ces valeurs $E(X_1 - Y_1)$ et $E((X_1 - Y_1)^2)$ par un calcul direct.
i) Soit $n \geq 1$. Calculer la fonction caractéristique φ_n de la variable aléatoire Z_n définie par $Z_n = \frac{1}{\sqrt{n}}\sum_{i=1}^n (X_i - Y_i)$. **j)** Lorsque $n \to \infty$, montrer que φ_n converge vers la fonction caractéristique d'une variable aléatoire de loi classique. Que peut-on en déduire sur le comportement de Z_n quand $n \to \infty$?
k) Retrouver ce résultat à l'aide d'un résultat du cours que l'on citera précisément.

CORRIGÉ a) Une loi exponentielle $\mathcal{E}(1)$ a pour densité $f_X : x \mapsto \mathbf{1}_{\mathbb{R}_+}(z)e^{-z}$. Graphe facile...
b) La fonction de répartition vaut $F_X(x) = \mathbf{1}_{\mathbb{R}_+}(x)(1 - e^{-x})$.
c) La fonction caractéristique vaut $\varphi_X(t) = E(e^{itX}) = \frac{1}{1-it}$.
d) On a $f_Z(x) = \frac{1}{\sqrt{2\pi}}e^{-x^2/2}$, et $\varphi_Z(t) = e^{-t^2/2}$.
e) On a ici $f_Z(z) = \frac{1}{\sqrt{2\pi}\sigma}e^{-(x-m)^2/(2\sigma^2)}$, et $\varphi_Z(t) = e^{itm-\sigma^2t^2/2}$.
f) Comme X_1 et Y_1 sont indépendantes, on a $\varphi_{X_1-Y_1}(t) = \varphi_{X_1}(t)\varphi_{-Y_1}(t) = \varphi_{X_1}(t)\varphi_{Y_1}(-t) = \frac{1}{1-it}\frac{1}{1+it} = \frac{1}{1+t^2}$.
g) Au voisinage de 0 on a $\varphi(t) = \varphi_{X_1-Y_1}(t) = 1 - t^2 + o(t^3)$. D'après le cours et la formule de Taylor en 0 on sait que $\varphi_{X_1-Y_1}(t) = 1 + iE(X_1 - Y_1)t - \frac{E((X_1-Y_1)^2)}{2}t^2 + o(t^2)$. On en déduit immédiatement $E(X_1 - Y_1) = 0$ et $E((X_1 - Y_1)^2) = 2$.
h) X_1 et Y_1 étant de même loi, on a par linéarité de l'espérance, $E(X_1 - Y_1) = E(X_1) - E(Y_1) = 1 - 1 = 0$. D'autre part, l'indépendance de X_1 et Y_1 fournit $E((X_1 - Y_1)^2) = E(X_1^2 + Y_1^2 - 2X_1Y_1) = E(X_1^2) + E(Y_1^2) - 2E(X_1)E(Y_1)$. Un calcul donne $E(X_1^2) = 2$, et donc $E((X_1 - Y_1)^2) = 2$.
i) On a $\varphi_n(t) = \varphi_{\sum_{i=1}^n (X_i - Y_i)}(\frac{t}{\sqrt{n}}) = \Pi_{i=1}^n \varphi_{X_i - Y_i}(\frac{t}{\sqrt{n}}) = \left(\varphi(\frac{t}{\sqrt{n}})\right)^n = \left(\frac{1}{1+t^2/n}\right)^n$
j) Quand $n \to \infty$ on écrit $\left(\frac{1}{1+t^2/n}\right)^n = (1 - \frac{t^2}{n}(1 + \varepsilon(\frac{t^2}{n}))^n = \exp(n \ln(1 - \frac{t^2}{n}(1 + \varepsilon(\frac{t^2}{n}))))$, où $\varepsilon(\frac{t^2}{n})$ est une fonction tendant vers 0 en 0. On en déduit que le terme dans l'exponentielle tend vers $-t^2$ quand $n \to \infty$ et donc que $\varphi_n(t) \to e^{-t^2}$. On reconnaît que la limite est la fonction caractéristique d'une variable aléatoire de loi normale $\mathcal{N}(0,2)$ de moyenne 0 et de variance $\sigma^2 = 2$. Ceci équivaut au fait que U_n converge en loi vers une variable aléatoire de loi normale $\mathcal{N}(0,2)$.
k) Le théorème limite central (voir cours) s'applique ici et donne exactement le résultat voulu.