

Contrôle Continu 1.

La qualité et la rigueur de la rédaction seront prises en compte. Vous indiquerez dans vos réponses, même brièvement, chaque usage d'une hypothèse de l'exercice ou d'un théorème du cours.

Exercice 1. Soient X, Y et U trois variables aléatoires qui satisfont les hypothèses suivantes :

- (H1) U est une variable de Bernoulli de paramètre $0 < p < 1$.
- (H2) X et Y sont des variables aléatoires réelles intégrables de densités f_X et f_Y par rapport à la mesure de Lebesgue. Nous noterons $m_X = \mathbb{E}[X]$ et $m_Y = \mathbb{E}[Y]$.
- (H3) (X, U) sont indépendantes et (Y, U) sont indépendantes.

Soit Z la v.a. réelle définie par $Z = X\mathbf{1}_{U=0} + Y\mathbf{1}_{U=1}$.

- a) Montrer que Z admet pour densité $f_Z(z) = (1 - p)f_X(z) + pf_Y(z)$.
- b) Calculer $\mathbb{E}[Z|U]$.
- c) Montrer que, presque sûrement,

$$\mathbb{E}[U|Z] = \frac{pf_Y(Z)}{(1 - p)f_X(Z) + pf_Y(Z)}$$

- d) Calculer $\mathbb{E}[X|Z]$ avec l'hypothèse supplémentaire que les (X, Y, U) sont toutes trois indépendantes.

Solution de l'exercice 1. Il s'agit ici de bien comprendre comment est construite la variable Z . On a un jeu biaisé de pile ou face avec la variable U : en cas de succès, on choisit pour Z la valeur X et, en cas d'échec, on choisit pour Z la valeur de Z . Toute expression utilisant Z est obtenue en choisissant la même expression sur X ou sur Y selon la valeur de U .

- a) **Méthode 1 :** (cette méthode a l'avantage de préparer la méthode de la fonction muette plus tard) Soit $g : \mathbb{R} \rightarrow \mathbb{R}$ mesurable bornée quelconque. Par une rapide étude de cas, nous avons

$$g(Z) = g(X)\mathbf{1}_{U=0} + g(Y)\mathbf{1}_{U=1}$$

Nous avons alors

$$\mathbb{E}[g(Z)] \stackrel{(\text{lin})}{=} \mathbb{E}[g(X)\mathbf{1}_{U=0}] + \mathbb{E}[g(Y)\mathbf{1}_{U=1}] \stackrel{(\text{H3})}{=} \mathbb{E}[g(X)]\mathbb{P}(U = 0) + \mathbb{E}[g(Y)]\mathbb{P}(U = 1)$$

Puis en utilisant (H2) et (H1), nous obtenons la formule intégrale

$$\mathbb{E}[g(Z)] = \int_{\mathbb{R}} g(z)(f_X(z)(1 - p) + f_Y(z)p)dz$$

et nous identifions la densité voulue de Z dans l'intégrande, à côté de g .

Méthode 2 On a l'union disjointe suivante

$$\{Z \leq t\} = \{X \leq t, U = 0\} \cup \{Y \leq t, U = 1\}$$

On a alors, en utilisant (H3) puis (H1) sur U , pour les fonctions de répartitions

$$F_Z(t) = (1 - p)F_X(t) + pF_Y(t)$$

On insère alors les densités selon (H2) dans le membre de droite et on récupère ainsi que $F_Z(t) = \int_{-\infty}^t f_Z(s)ds$ avec la densité voulue.

- b) Tout d'abord, $|Z| \leq |X| + |Y|$ par majoration des indicatrices par 1 et inégalité triangulaire. Donc Z est intégrable. De plus, la linéarité de l'espé. condit. donne

$$\mathbb{E}[Z|U] = \mathbb{E}[X\mathbf{1}_{U=0}|U] + \mathbb{E}[Y\mathbf{1}_{U=1}|U]$$

Les événements $U = 0$ et $U = 1$ sont $\sigma(U)$ -mesurables donc

$$\mathbb{E}[Z|U] = \mathbf{1}_{U=0}\mathbb{E}[X|U] + \mathbf{1}_{U=1}\mathbb{E}[Y|U]$$

Les indépendances respectives de X et Y par rapport à U (H3) donnent alors

$$\mathbb{E}[Z|U] = \mathbf{1}_{U=0}\mathbb{E}[X] + \mathbf{1}_{U=1}\mathbb{E}[Y] = m_X\mathbf{1}_{U=0} + m_Y\mathbf{1}_{U=1}$$

où l'on utilisé (H2) pour l'identification des espérance.

- c) Soit $g : \mathbb{R} \rightarrow \mathbb{R}$ mesurable bornée arbitraire. Nous observons tout d'abord que $g(Z) = g(X\mathbf{1}_{U=0} + Y\mathbf{1}_{U=1}) = \mathbf{1}_{U=0}g(X) + \mathbf{1}_{U=1}g(Y)$ par simple étude de cas. Nous avons alors :

$$\mathbb{E}[Ug(Z)] \stackrel{(\text{lin})}{=} \mathbb{E}[U\mathbf{1}_{U=0}g(X)] + \mathbb{E}[U\mathbf{1}_{U=1}g(Y)] = 0 + \mathbb{E}[\mathbf{1}_{U=1}g(Y)] = p\mathbb{E}[g(Y)] = p \sum_{n \in \mathbb{N}} \beta(n)g(n)$$

par indépendance (H3) et formule de transfert sur U et Y (H1 et H2)

Par ailleurs, pour toute fonction $\phi : \mathbb{N} \rightarrow \mathbb{R}$ mesurable bornée, nous avons par formule de transfert sur la loi de Z (question (a)) :

$$\mathbb{E}[\phi(Z)g(Z)] = \int \phi(z)g(z) ((1-p)f_X(z) + pf_Y(z))$$

En prenant $h(z) = pf_Y(z)/((1-p)f_X(z) + pf_Y(z))$ pour tout z tel que $f_Z(z) \neq 0$ et n'importe quelle autre valeur ailleurs, nous avons $\mathbb{E}[Ug(Z)] = \mathbb{E}[h(Z)g(Z)]$ et, par la caractérisation du cours,

$$\mathbb{E}[U]Z = h(Z) \quad \text{p.s.}$$

- d) Soit $g : \mathbb{N} \rightarrow \mathbb{R}$ mesurable bornée arbitraire. Nous observons tout d'abord que $g(Z) = g(X\mathbf{1}_{U=0} + Y\mathbf{1}_{U=1}) = \mathbf{1}_{U=0}g(X) + \mathbf{1}_{U=1}g(Y)$ par simple étude de cas. Nous avons alors :

$$\begin{aligned} \mathbb{E}[Xg(Z)] &= \mathbb{E}[X\mathbf{1}_{U=0}g(X)] + \mathbb{E}[X\mathbf{1}_{U=1}g(Y)] = (1-p)\mathbb{E}[Xg(X)] + p\mathbb{E}[X]\mathbb{E}[g(Y)] \\ &= \int_{\mathbb{R}} g(z) [(1-p)zf_X(z) + pm_Xf_Y(z)] \end{aligned}$$

par indépendance (hypo supplémentaire) et formules de transfert sur X et Y respectivement. En posant

$$h(z) = \frac{(1-p)zf_X(z) + pm_Xf_Y(z)}{(1-p)f_X(z) + pf_Y(z)}$$

si $f_Z(z) \neq 0$ et n'importe quelle valeur ailleurs, nous avons $\mathbb{E}[Ug(Z)] = \mathbb{E}[h(Z)g(Z)]$ et, par la caractérisation du cours,

$$\mathbb{E}[U]Z = \frac{(1-p)Z\alpha(Z) + pm_X\beta(Z)}{(1-p)\alpha(Z) + p\beta(Z)} \quad \text{p.s.}$$

Commentaire après correction.

En finir avec les calculs de lois. Pour une variable entière ou réelle ou abstraite, il s'agit de calculer $\mathbb{E}[g(Z)]$ pour toute fonction f mesurable bornée, et après simplification, aboutir à

$$\mathbb{E}[g(Z)] = \int_E g(x)d\mu_Z(x)$$

Pour une variable entière, cette intégrale devient une somme. Si $Z = (X, Y)$ est un vecteur aléatoire de \mathbb{R}^2 à densité, c'est une intégrable double, etc. Calculer $\mathbb{E}[Z]$ ne suffit pas : par exemple $\text{Unif}([-1, 1])$ et $\mathcal{N}(0, 1)$ ont la même espérance mais pas du tout la même loi.

Il est parfois utile de considérer d'autres caractérisations. Par exemple, pour Z entière, on peut calculer directement $\mathbb{P}(Z = n)$ pour tout $n \in \mathbb{N}$. Pour Z réelle, on peut alors calculer $F_Z(t) = \mathbb{P}(Z \leq t)$ pour tout $t \in \mathbb{R}$ (attention, F_Z n'est pas toujours dérivable...).

En finir avec les discussions de cas. L'expérience montre qu'on aime souvent raisonner avec des études de cas, par exemple, dans le premier exercice ici, écrire

$$Z = \begin{cases} X & \text{si } U = 0 \\ Y & \text{si } U = 1 \end{cases}$$

C'est utile pour comprendre ce qui se passe intuitivement mais est d'une utilité quasi-nulle dans les calculs, en particulier car ω apparaît partout (dans Z, X, Y et U). Aucune formule connue ne manipule ce genre d'objets. En revanche, il existe une alternative rigoureuse équivalente particulièrement adaptée aux probabilités : l'usage de partitions et de sommes d'indicatrices.

Si on a, pour des événements $(A_k)_{1 \leq k \leq n}$ en nombre fini ou dénombrable qui forment une partition de Ω (au pire on ajoute par le complémentaire de leur union...),

$$Z(\omega) = \begin{cases} X_1(\omega) & \text{si } \omega \in A_1 \\ X_2(\omega) & \text{si } \omega \in A_2 \\ \dots & \\ X_n & \text{si } \omega \in A_n \end{cases}$$

La bonne représentation pour les calculs sur Z ou un événement B quelconque est la suivante :

$$Z = \sum_{k=1}^n X_k \mathbf{1}_{A_k} \qquad B = \bigcup_{k=1}^n (B \cap A_k)$$

(union disjointe ici). Cela permet d'utiliser à la fois la σ -additivité ou encore la linéarité de l'espérance (et à partir de maintenant l'espérance conditionnelle $\mathbb{E}[\bullet | \sigma(A_k; 1 \leq k \leq n)]$). Vous pouvez également vous convaincre (faites-le rigoureusement au moins une fois dans votre cursus) que, pour $h : \mathbb{R} \rightarrow \mathbb{R}$, on a aussi $h(Z) = \sum_{k=1}^n h(X_k) \mathbf{1}_{A_k}$ et cela permet souvent d'avancer rigoureusement dans les calculs.

Cessez de vous entêter avec des cas alors qu'un outil parfaitement adéquat est disponible !

Attention au piège avec les cas (c'est pour ça qu'il ne faut pas les utiliser) Contrairement à ce que j'ai vu dans beaucoup de copies, $\mathbb{P}(Z = X)$ n'est pas nécessairement égal $1 - p$ ici ! En effet, considérez par exemple $X = Y$ (cela respecte toutes les hypothèses !) et vous aurez $\mathbb{P}(Z = X) = 1$: la variable Y peut être égale à X pour certains ω ! Par exemple, ici il suffisait d'écrire la partition sur U :

$$\begin{aligned} \{Z = X\} &= \{Z = X, U = 0\} \cup \{Z = X, U = 1\} = \{X = X, U = 0\} \cup \{Y = X, U = 1\} \\ &= \{U = 0\} \cup \{Y = X, U = 1\} \end{aligned}$$

(union disjointe) pour voir tout de suite que

$$\mathbb{P}(Z = X) = (1 - p) + p\mathbb{P}(Y = X)$$

car U est indépendante de Y et X (et donc de $Y = X$ même si Y et X ne sont pas indépendantes : exercice !).

Exercice 2. Soit $f : \mathbb{R} \rightarrow \mathbb{R}$ une fonction K -lipshitzienne, i.e. telle que, pour tout $x, y \in \mathbb{R}$, $|f(x) - f(y)| \leq K|x - y|$. Nous supposons dans tout l'exercice que $K < 1$.

Rappel : un résultat classique énonce qu'une suite (déterministe) $x_{n+1} = f(x_n)$ converge vers l'unique point fixe l de f (qui existe bien). Par ailleurs, pour tout $q > K$, on a $|x_n - l| = o(q^n)$.

Nous nous intéressons à une version avec aléa de ce résultat. Soit $(U_n)_{n \in \mathbb{N}}$ une suite de v.a. intégrables, avec $\mathbb{E}[|U_n|] = \alpha_n < \infty$. Soit $x_0 \in \mathbb{R}$ un réel fixé. Soit $(X_n)_{n \in \mathbb{N}}$ la suite de v.a. définie par la récurrence $X_0 = x_0$ et $X_{n+1} = f(X_n + U_{n+1})$.

- a) Montrer que, pour tout $n \in \mathbb{N}$, $\mathbb{E}[|X_{n+1} - l|] \leq K\mathbb{E}[|X_n - l|] + K\alpha_{n+1}$.
- b) Montrer que, pour tout $n \in \mathbb{N}$,

$$\mathbb{E}[|X_n - l|] \leq K^n|x_0 - l| + \sum_{k=1}^n \alpha_k K^{n+1-k}$$

- c) Montrer que $\sum_{n \geq 0} \mathbb{E}[|X_n - l|] \leq C_0 + C_1 \sum_{k \geq 1} \alpha_k$ pour des constantes C_0 et C_1 , dont on donnera une expression possible.
- d) Montrer que si $\sum_{k \geq 1} \alpha_k < \infty$, alors (X_n) converge presque sûrement vers l .
- e) *Bonus* : on dit que « $A_n = o_{\text{p.s.}}(B_n)$ » si et seulement si $A_n/B_n \xrightarrow{\text{p.s.}} 0$. Montrer que, pour tout $q > \max(K, 1/R)$, $|X_n - l| = o_{\text{p.s.}}(q^n)$, où R est le rayon de convergence de la série entière $\sum_k \alpha_k z^k$. Comparez au résultat déterministe.

Solution de l'exercice 2. Le but de cet exercice est de comprendre l'influence de l'insertion de bruit aléatoire dans un exercice classique de première année de mathématique (convergence de (x_n) vers le point fixe l de f et sa vitesse de convergence). Nous voyons ici que la convergence n'est pas modifiée (question (d)). Si le bruit est faible, la vitesse de convergence ne bouge pas mais, si le bruit est trop fort, la convergence peut être ralentie (question (e)).

- a) On a $|X_{n+1} - l| = |f(X_n + U_{n+1}) - f(l)| \leq K|X_n + U_{n+1} - l|$ car f est K -lipshitzienne. Par inégalité triangulaire, nous avons ensuite $|X_{n+1} - l| \leq K|X_n - l| + |U_{n+1}|$ p.s. Par monotonie de l'espérance positive, nous avons alors $\mathbb{E}[|X_{n+1} - l|] \leq K\mathbb{E}[|X_n - l|] + K\alpha_{n+1}$.
- b) Cela se montre directement par récurrence. Pour $n = 0$, nous avons $\mathbb{E}[|X_0 - l|] = |x_0 - l|$ par initialisation de X_0 et le résultat est vrai. Supposons le résultat vrai pour un n donné. Alors l'utilisation de la question précédente donne :

$$\begin{aligned} \mathbb{E}[|X_{n+1} - l|] &\leq K\mathbb{E}[|X_n - l|] + K\alpha_{n+1} \leq K \left(K^n|x_0 - l| + \sum_{k=1}^n \alpha_k K^{n+1-k} \right) + K\alpha_{n+1} \\ &\leq K^{n+1}|x_0 - l| + \sum_{k=1}^n \alpha_k K^{n+2-k} + K\alpha_{n+1} \end{aligned}$$

qui est le résultat attendu en regroupant le dernier terme dans la somme pour $k = n + 1$.

- c) Un calcul direct donne

$$\begin{aligned} \sum_{n \geq 0} \mathbb{E}[|X_n - l|] &\leq |x_0 - l| \sum_{n \geq 0} K^n + \sum_{n \geq 0} \sum_{k=1}^n \alpha_k K^{n+1-k} \leq \frac{|x_0 - l|}{1 - K} + \sum_{k \geq 1} \sum_{n \geq k} \alpha_k K^{n+1-k} \\ &\leq \frac{|x_0 - l|}{1 - K} + \sum_{k \geq 1} \alpha_k \sum_{n' \geq 0} K^{n'+1} \leq \frac{|x_0 - l|}{1 - K} + \sum_{k \geq 1} \alpha_k \frac{K}{1 - K} \end{aligned}$$

en utilisant $K < 1$ pour la convergence de la série géométrique et posant $n' = n - k$. On peut prendre $C_0 = |x_0 - l|/(1 - K)$ et $C_1 = K/(1 - K)$.

- d) Deux méthodes sont possibles. **Méthode 1** : par convergence monotone, on observe que $\mathbb{E}[\sum_{n \geq 0} |X_n - l|] = \sum_{n \geq 0} \mathbb{E}[|X_n - l|] < \infty$ d'après l'hypothèse sur la série des α_k et la question précédente. La v.a. $S = \sum_{n \geq 0} |X_n - l|$ est donc finie p.s. car d'espérance finie donc le terme général $|X_n - l|$ tend p.s. vers 0.

Méthode 2 : par l'inégalité de Markov, on a pour tout $p \in \mathbb{N}$,

$$\sum_{n \geq 0} \mathbb{P}(|X_n - l| \geq 2^{-p}) \leq 2^p \sum_{n \geq 0} \mathbb{E}[|X_n - l|] < \infty$$

On obtient d'une part la convergence en probabilité mais surtout, par le lemme de Borel-Cantelli, $\mathbb{P}(\limsup_n \{|X_n - l| \geq 2^{-p}\}) = 0$ et donc $\mathbb{P}(\liminf_n \{|X_n - l| < 2^{-p}\}) = 1$ puis par intersection d'événements décroissants, $\mathbb{P}(\bigcap_{p \geq 0} \liminf_n \{|X_n - l| < 2^{-p}\}) = 1$, qui est la définition de la convergence presque sûre.

e) C'est une simple amélioration des questions précédentes. Soit $q > \max(K, 1/R)$. On a

$$\sum_{n \geq 0} \mathbb{E}[q^{-n}|X_n - l|] \leq |x_0 - l| \sum_{n \geq 0} (K/q)^n + q \left(\sum_{k \geq 1} \alpha_k q^{-k} \right) \sum_{n' \geq 0} (K/q)^{n'+1}$$

L'hypothèse $q > K$ donne la convergence des séries géométriques en K/q . L'hypothèse $q > 1/R$ donne $1/q < R$ et la série $\sum_{k \geq 1} \alpha_k q^{-k}$ est également convergente. Le même raisonnement qu'à la question précédente donne alors $|X_n - l|/q^n \rightarrow 0$ p.s., i.e. $|X_n - l| = o(q^n)$ p.s.

L'interprétation est la suivante : si l'aléa est suffisamment faible pour n grand, i.e. α_k décroît suffisamment vite pour que $R > K$, alors la vitesse de convergence en q^n n'est pas affectée par rapport au cas déterministe. Si l'aléa est trop fort, i.e. $1/R > K$, alors il est le principal frein à la convergence en q^n .

Commentaire après correction.

Du rôle des formules qu'on vous donne. *Il arrive qu'on vous donne une formule à démontrer (cf. I.a, I.c, II.a, II.b). Elles sont là pour vous permettre d'avancer même si vous n'arriveriez pas à les trouver seuls et pour vous éviter des erreurs de calcul en cascade. Mais il s'agit de les démontrer rigoureusement. Les recopier en énonçant seulement et rapidement un pseudo-argument ou quelques mots-clefs que vous imaginez à partir de la formule ne rapporte aucun point et vous induit souvent en erreur pour la suite.*

Du point de vue de votre réponse, il faut imaginer qu'on ne vous donne pas le résultat et que vous deviez le trouver et le justifier entièrement et rigoureusement. Vous le donner n'est pas prémâcher le travail...

Récurrence *La question II.b se montrait par récurrence. Soyons sérieux : quand on arrive en M1 de maths, on doit être capable de mettre en place en 30 secondes une récurrence propre. Écrire dans une copie, "le résultat de (II.b) se montre par récurrence" et/ou itérer le résultat de (I.a) de n à $n - 2$ ou $n - 3$ dans un coin de feuille et dire "ça marche" ne rapporte aucun point.*

Borel-Cantelli *Prononcez "Borel-Cantelli" comme une formule magique issue de Harry Potter et dire "ça converge p.s." ne rapporte aucun point. Il s'agit au minimum : (1) d'écrire rigoureusement les événements auxquels on l'applique (2) de prouver rigoureusement l'hypothèse de base du lemme de Borel-Cantelli de convergence de la série de probabilités, (3) d'écrire rigoureusement la conclusion du lemme de Borel-Cantelli et de la relier, même avec des mots, avec le résultat à montrer.*