CHAPITRE

14 Aléatoire

Dès son plus jeune âge, Florence NIGHTINGALE se révèle particulièrement douée pour les mathématiques et y excelle grâce aux enseignements de son père. Elle est aussi l'élève du mathématicien James Joseph SYLVESTER. Elle s'intéresse notamment à la statistique, un domaine dans lequel son père, un des pionniers de l'épidémiologie, est un expert. Elle a fréquemment recours aux analyses statistiques dans ses compilations, analyses et présentations de données sur les soins médicaux et la santé publique.

Florence NIGHTINGALE est une pionnière de la présentation visuelle de l'information. Elle utilise fréquemment les coxcombs pour présenter des rapports sur la nature et l'ampleur des conditions des soins médicaux pendant la Guerre de Crimée aux membres du Parlement et aux fonctionnaires, qui n'auraient probablement pas pu lire ou comprendre des rapports statistiques traditionnels.

Le principe dans la bonne société est que les malades sont soignés chez eux ; les hôpitaux sont pour les pauvres. Florence NIGHTINGALE prouva que 90% des patients des hôpitaux de Londres trouvaient la mort contre seulement 60% pour les malades ne se rendant pas à l'hôpital.

En 1857, d'après la BBC, elle est la femme la plus célèbre après la Reine Victoria elle-même. En 1858, elle fut la première femme à être élue membre de la Royal Statistical Society, et devient par la suite membre honoraire de l'American Statistical Association.

Introduction

Programme

— Couple de variables aléatoires. Loi conjointe, lois marginales. Loi conditionnelle de Y sachant $X = x$, $X > x$ ou autres inégalités.

— Extension aux n-uplets de variables aléatoires. Vecteurs aléatoires discrets.

— Couple et famille de variables aléatoires mutuellement indépendantes.

— Si $X_1, X_2, \ldots, X_n$ sont des variables aléatoires mutuellement indépendantes, alors pour tout m compris entre 1 et $n-1$, et toutes fonctions f et g, les variables $f(X_1, \ldots, X_m)$ et $g(X_{m+1}, \ldots, X_n)$ sont indépendantes (démonstration non exigible).

— Existence d'espaces probabilisés portant une suite de variables indépendantes de lois discrètes données (démonstration hors programme). Modélisation du jeu de pile ou face infini.

— Si X et Y sont deux variables aléatoires indépendantes d'espérances finies, alors : $\mathbf{E}(XY) = \mathbf{E}(X) \cdot \mathbf{E}(Y)$.

— Covariance. Relation $\mathrm{Cov}(X, Y) = \mathbf{E}(XY) - \mathbf{E}(X)\mathbf{E}(Y)$. Cas de variables indépendantes. Variance d'une somme finie, cas de variables deux à deux indépendantes.

— Loi faible des grands nombres. Les étudiant(e)s doivent savoir retrouver, pour $\varepsilon > 0$, l'inégalité $P\left(\left|\dfrac{S_n}{n} - m\right| \geq \varepsilon\right) \leq \dfrac{\sigma^2}{n\varepsilon^2}$.

— Fonction génératrice d'une somme finie de variables aléatoires indépendantes à valeurs dans $\mathbf{N}$.

Il y a de nombreux modes de convergence que ce soit en analyse ou en probabilités. Comme une variable aléatoire est une fonction, définie sur Ω, on peut penser aux modes de convergence des fonctions, mais c'est en général inadapté. Si $(X_n)_{n \in \mathbf{N}^*}$ est une suite de variables aléatoires, alors X_n converge simplement vers X veut dire que pour tout événement aléatoire ω, on a $\lim X_n(\omega) = X(\omega)$. Par exemple si X est une constante (la variable aléatoire identiquement égale à une certaine constante) et si X_n est la proportion de Pile qui a été obtenue lors du tirage ω, un tel phénomène ne se produira pas : il existe des ω qui ne donnent que des Pile, ou que des Face.

Un mode moins exigeant est la convergence presque sûre.

Définition 14 - 1

Convergence presque sûre ♠

On dit que X_n converge presque sûrement vers X si l'ensemble des ω dans Ω tels qu'on ait $\lim X_n(\omega) = X(\omega)$ est un événement presque sûr. Autrement dit si $\mathbf{P}(\lim X_n = X) = 1$. On note alors $X_n \xrightarrow{ps} X$.

Remarque 14 - 1

Il revient au même de demander que, pour tout ε dans $\mathbf{R}_+^*$, l'ensemble des ω dans Ω tels que $|X_n - X|$ soit supérieur à ε pour une infinité d'entiers n, soit négligeable.

C'est en particulier le cas si la série $\sum \mathbf{P}(|X_n - X| > \varepsilon)$ converge. En effet, notons

$$A_n = (|X_n - X| > \varepsilon) \quad \text{et} \quad B_n = \bigcup_{k=n}^{+\infty} A_k$$

de sorte que l'événement « A_n a lieu une infinité de fois » est l'événement

$$\bigcap_{k=1}^{+\infty} B_k \; .$$

Cet événement a donc, par croissance, une probabilité inférieure à $\mathbf{P}(B_n)$ pour tout entier n et donc aussi à $\displaystyle\sum_{k=n}^{+\infty} \mathbf{P}(A_k)$ par sous-additivité. Si la série $\sum \mathbf{P}(|X_n - X| > \varepsilon)$, converge il s'agit là du reste d'une série convergente et donc c'est arbitrairement petit. On en déduit qu'on a affaire à un événement de probabilité nulle, i.e. que X_n converge presque sûrement vers X.

Proposition 14 - 1

Critère de convergence presque sûre ♠
Si la série $\sum \mathbf{P}(|X_n - X| > \varepsilon)$ converge, alors X_n converge presque sûrement vers X.

La convergence presque sûre est un résultat souvent technique difficile à obtenir. Moins contraignant est la convergence en probabilité.

Définition 14 - 2

On dit qu'une suite $(X_n)_{n\in\mathbf{N}^*}$ de variables aléatoires définies sur le même espace probabilisé $(\Omega, \mathcal{A}, \mathbf{P})$ converge en probabilité vers une variable aléatoire X, également définie sur $(\Omega, \mathcal{A}, \mathbf{P})$, si

$$\forall \varepsilon \in \mathbf{R}_+^* \qquad \lim_{n\to+\infty} \mathbf{P}(|X_n - X| > \varepsilon) = 0 \; .$$

On note $X_n \xrightarrow{\mathbf{P}} X$.

Remarque 14 - 2

Par définition la convergence presque sûre entraîne la convergence en probabilité, ce qui est cohérent avec le fait que le terme général d'une série convergente tend vers 0,

Exemple 14 - 1

Soit (X_n) une suite de variables aléatoires ayant un moment d'ordre 2 et telles que $\mathbf{E}(X_n) = 0$ pour tout n, et $\lim \mathrm{V}(X_n) = 0$. Alors l'inégalité de Bienaymé-Tchebychev entraîne que X_n converge vers 0 en probabilité.

Propriété 14 - 1

Si $X_n \xrightarrow{\mathbf{P}} X$ et $Y_n \xrightarrow{\mathbf{P}} Y$, alors pour tous scalaires α et β, $\alpha X_n + \beta Y_n \xrightarrow{\mathbf{P}} \alpha X + \beta Y$.

Aparté

On parle de convergence en loi, lorsque la loi de X_n converge vers celle de X. On dispose encore d'autres modes de convergence. Citons la convergence en moyenne, ou dans L^1, i.e. $\lim \mathbf{E}(|X_n - X|) = 0$.

Plus généralement, on parle de convergence dans L^p (pour $p \geq 1$), si $\lim \mathbf{E}(|X_n - X|^p) = 0$.

1 Familles de variables aléatoires

Dans l'étude des phénomènes aléatoires, notamment par répétition d'expérience, on a besoin de modéliser des familles de variables aléatoires. Dans ce cas les variables aléatoires sont souvent indépendantes les unes des autres, mais pas toujours comme dans le cas d'un tirage sans remise. On peut également s'intéresser à des variables aléatoires à valeurs vectorielles et donc à des n-uplets de variables aléatoires.

Dans cette section X et Y sont des variables aléatoires discrètes définies sur $(\Omega, \mathcal{A}, \mathbf{P})$ et à valeurs dans E.

Puisque le couple (X, Y) est à valeurs dans $X(\Omega) \times Y(\Omega)$ et que le produit de deux ensembles au plus dénombrables est au plus dénombrable, (X, Y) est une variable aléatoire discrète.

Définition 14 - 3

Le couple (X, Y) est une variable aléatoire discrète à valeurs dans E^2. On appelle loi conjointe la loi de (X, Y).
Les variables X et Y sont appelées marginales du couple (X, Y) et leurs lois sont appelées lois marginales (en X ou en Y) du couple (X, Y).

Danger

Les lois marginales ne déterminent pas la loi conjointe. En effet on s'intéresse aux événements $(X(\omega), Y(\omega))$ et ceux-ci peuvent, par exemple, ne pas décrire $X(\Omega) \times Y(\Omega)$.
Prenons l'exemple de deux dés à six faces. Les lois marginales sont des lois uniformes sur $[\![1;6]\!]$. Si on lance les dés indépendamment, la loi conjointe est uniforme sur le carré $[\![1;6]\!]^2$. Si au contraire on prend pour second dé le premier, on obtient une loi uniforme sur le segment diagonal du carré précédent. On pourrait aussi prendre la face opposée comme second dé, ce qui conduirait à une autre diagonale du carré.

Puisqu'on a affaire à une variable discrète, connaître la loi de (X, Y), c'est connaître les probabilités atomiques $p_{k,\ell}$ avec $X(\Omega) = \{x_k \mid k \in K\}$, $Y(\Omega) = \{y_\ell \mid \ell \in L\}$ et $p_{k,\ell} = \mathbf{P}(X = x_k, Y = y_\ell)$, où K et L sont au plus dénombrables. Avec ces notations, il est à noter qu'on peut très bien avoir $p_{k,\ell} = 0$ pour certains couples (k, ℓ). La loi des probabilités totales donne

$$\sum_{(k,\ell)\in K\times L} p_{k,\ell} = 1\,, \quad \mathbf{P}(X = x_k) = \sum_{\ell\in L} p_{k,l} \quad \text{et} \quad \mathbf{P}(Y = y_\ell) = \sum_{k\in K} p_{k,\ell}\,.$$

Définition 14 - 4

Soit y dans $Y(\Omega)$ tel que $(Y = y)$ ne soit pas négligeable. On appelle loi conditionnelle de X sachant $(Y = y)$, la loi de X dans l'espace probabilisé $(\Omega, \mathcal{A}, \mathbf{P}_{(Y=y)})$. On note $\mathbf{P}\,(X = x \mid Y = y)$ la quantité $\mathbf{P}_{(Y=y)}\,(X = x)$. Autrement dit

$$\mathbf{P}\,(X = x \mid Y = y) = \frac{\mathbf{P}(X = x, Y = y)}{\mathbf{P}(Y = y)} = \frac{p_{k,\ell}}{\sum_{j\in K} p_{j,\ell}}\,.$$

Définition 14 - 5

On dit que X et Y sont indépendants si, pour tous x dans $X(\Omega)$ et y dans $Y(\Omega)$, les événements $(X = x)$ et $(Y = y)$ sont indépendants, i.e.

$$\mathbf{P}(X = x, Y = y) = \mathbf{P}(X = x) \times \mathbf{P}(Y = y) \ .$$

En particulier, pour tout événement non négligeable $(Y = y)$, la loi conditionnelle de X sachant $(Y = y)$ est égale à la loi de X.

Autrement dit le fait que X prenne une valeur donnée est indépendant du fait que Y prenne telle ou telle valeur.

Remarque 14 - 3

Si X et Y sont indépendants et si A et B sont des parties de E, on a

$$\mathbf{P}\left((X, Y) \in A \times B\right) = \mathbf{P}(X \in A) \times \mathbf{P}(Y \in B) \ .$$

En effet la famille $(\mathbf{P}(X = x, Y = y))_{(x,y)\in A\times B}$ est aussi, par indépendance, la famille $(\mathbf{P}(X = x)\mathbf{P}(Y = y))_{(x,y)\in A\times B}$. Les deux quantités étudiées en sont la somme commune en vertu du théorème de sommations par paquets pour les familles à termes positifs.

Exemple 14 - 2

On se donne un événement rare qui n'a que deux issues possibles et on cherche combien de fois une issue donnée a lieu. On modélise le nombre d'événements par une loi de Poisson et chacun des événements donne lieu à un événement aléatoire suivant une loi de Bernoulli.

On note N la variable aléatoire correspondant au nombre d'événements, et donc $N \sim \mathcal{P}(\lambda)$. On note X la variable aléatoire correspondant au nombre d'événements ayant eu pour issue une des deux issues. On pourrait écrire $X = \sum_{k=1}^{N} X_i$ avec $X_i \sim \mathcal{B}(p)$, mais il faut prendre garde que N est une variable aléatoire ! On a donc

$$X(\omega) = \sum_{k=1}^{N(\omega)} X_i(\omega) \ .$$

Pour étudier cet exemple de façon élégante, on peut utiliser la notion, hors-programme, d'espérance conditionnelle. On va se contenter de lois conditionnelles et montrer $X \sim \mathcal{P}(\lambda p)$. On détermine d'abord la loi conjointe de (N, X). On a

$$\mathbf{P}(N = n, X = k) = \mathbf{P}(N = n)\mathbf{P}\left(X = k \,|\, N = n\right) = e^{-\lambda}\frac{\lambda^n}{n!}\binom{n}{k}p^k(1-p)^{n-k}$$

puisque la loi conditionnelle de X sachant $N = n$ est une loi binomiale $\mathcal{B}(n, p)$. Bien entendu, si $k > n$, la probabilité précédente est nulle par convention. On en déduit la loi marginale selon X par la formule

$$\mathbf{P}(X = k) = \sum_{n=k}^{+\infty} e^{-\lambda}\frac{\lambda^n}{n!}\binom{n}{k}p^k(1-p)^{n-k}$$

et donc $\mathbf{P}(X = k) = e^{-\lambda}\dfrac{\lambda^k}{k!}p^k e^{\lambda(1-p)}$. Autrement dit $X \sim \mathcal{P}(\lambda p)$.

En particulier $\mathbf{E}(X) = \lambda p$, ce qui est conforme à l'intuition.

Proposition 14 - 2

Si X et Y sont indépendants et admettent chacun une espérance, alors XY aussi et on a $\mathbf{E}(XY) = \mathbf{E}(X)\mathbf{E}(Y)$.

Démonstration. Puisque X et Y admettent une espérance, les familles $(x\mathbf{P}(X = x))_{x\in X(\Omega)}$ et $(y\mathbf{P}(Y = y))_{y\in Y(\Omega)}$ sont sommables. D'après le théorème de sommation par paquets dans le cas positif, on en déduit que la famille $(xy\mathbf{P}(X = x)\mathbf{P}(Y = y))_{(x,y)\in X(\Omega)\times Y(\Omega)}$ est sommable. Par indépendance de X et Y, ceci signifie que XY admet une espérance. Le théorème de sommation par paquets dans le cas quelconque permet de conclure $\mathbf{E}(XY) = \mathbf{E}(X)\mathbf{E}(Y)$. □

Remarque 14 - 4

En général l'équation $\mathbf{E}(XY) = \mathbf{E}(X)\mathbf{E}(Y)$ est fausse. Par exemple, si $X = Y$, on ne peut avoir $\mathbf{E}(XY) = \mathbf{E}(X)\mathbf{E}(Y)$ que si $\mathrm{V}(X) = 0$, i.e. si X est presque sûrement constante. En particulier si X n'admet pas de variance mais admet une espérance, $\mathbf{E}(XY) = +\infty$ alors que $\mathbf{E}(X)\mathbf{E}(Y)$ est fini.

On peut aussi choisir X et Y de sorte que XY soit nul identiquement sans que X ni Y ne soit d'espérance nulle, par exemple avec X et Y des variables aléatoires positives.

Danger

Toutefois il peut arriver qu'on ait $\mathbf{E}(XY) = \mathbf{E}(X)\mathbf{E}(Y)$ sans que X et Y soient indépendants. Autrement dit la réciproque de la proposition précédente est fausse en général.

On choisit X vérifiant $X \sim \mathcal{U}(\llbracket -1;1 \rrbracket)$ et $Y = |X|$. On a donc $Y \sim \mathcal{B}\left(\frac{2}{3}\right)$, $XY = X$ et donc $\mathbf{E}(XY) = \mathbf{E}(X) = 0 = \mathbf{E}(X)\mathbf{E}(Y)$. Par contre $\mathbf{P}(X = 1, Y = 0) = 0$ tandis que $\mathbf{P}(X = 1)\mathbf{P}(Y = 0) = \frac{1}{9}$ et donc X et Y ne sont pas indépendants.

Remarque 14 - 5

On se méfiera tout de même des exemples trop simples. Par exemple pour deux variables suivant des lois de BERNOULLI, on a effectivement équivalence entre X et Y indépendants et $\mathbf{E}(XY) = \mathbf{E}(X)\mathbf{E}(Y)$.

Définition 14 - 6

Si X, Y et XY admettent tous les trois des espérances, on appelle covariance de X et Y la quantité définie par $\mathrm{Cov}(X,Y) = \mathbf{E}((X - \mathbf{E}(X))(Y - \mathbf{E}(Y)))$, i.e.

$$\boxed{\mathrm{Cov}(X,Y) = \mathbf{E}(XY) - \mathbf{E}(X)\mathbf{E}(Y)}\ .$$

En particulier, si X et Y sont indépendants, leur covariance est nulle.

Cette définition est licite puisque, par linéarité, si X et Y admettent une espérance, alors XY en admet une si et seulement si $(X - \mathbf{E}(X))(Y - \mathbf{E}(Y))$ en admet une. La seconde formule pour la covariance résulte de la linéarité de l'espérance.

Remarque 14 - 6

La covariance représente un produit scalaire ainsi qu'en témoigne l'inégalité de Cauchy-Schwarz. En centrant la représentation du couple (X, Y) en sa moyenne et en normalisant les axes, i.e. en supposant qu'on a affaire à des variables centrées réduites, une covariance proche de 1 donne naissance à un nuage de points le long de la première bissectrice, une covariance proche de -1 à un nuage de points le long de la seconde bissectrice. Une covariance proche de zéro correspond à un nuage de points bien réparti ou à des points le long des axes ou encore à un mélange entre ces deux situations.

On déduit directement de la définition les propriétés élémentaires de la covariance :

Propriétés 14 - 2

L'application $(X, Y) \mapsto \text{Cov}\,(X, Y)$, définie sur l'espace $L^2(\Omega, \mathcal{A}, \mathbf{P})$ des variables aléatoires admettant un moment d'ordre 2, est bilinéaire symétrique positive. Autrement dit

1. on a $\text{Cov}\,(X, Y) = \text{Cov}\,(Y, X)$;
2. on a, pour α et β scalaires, $\text{Cov}\,(\alpha X, \beta Y) = \alpha\beta \text{Cov}\,(X, Y)$;
3. on a, pour Z dans $L^2(\Omega, \mathcal{A}, \mathbf{P})$, $\text{Cov}\,(X + Y, Z) = \text{Cov}\,(X, Z) + \text{Cov}\,(Y, Z)$;
4. en particulier, pour α et β scalaires, on a $\text{Cov}\,(\alpha X + \beta, Y) = \alpha \text{Cov}\,(X, Y)$;
5. et enfin $\text{Cov}\,(X, X) = \text{V}\,(X) \geq 0$.

Pour aller plus loin

On en déduit également une reformulation de l'inégalité de Cauchy-Schwarz ainsi que son cas d'égalité : $|\text{Cov}\,(X, Y)| \leq \sigma\,(X)\,\sigma\,(Y)$ avec égalité si et seulement si $X - \mathbf{E}\,(X)$ et $Y - \mathbf{E}\,(Y)$ sont colinéaires.

On appelle coefficient de corrélation le réel compris entre -1 et 1, défini si X et Y admettent des variances, celles-ci étant strictement positives, et donné par

$$\rho(X, Y) = \frac{\text{Cov}\,(X, Y)}{\sigma\,(X)\,\sigma\,(Y)}\,.$$

Remarque 14 - 7

Égalité de Bienaymé

Si X et Y sont indépendants et admettent une variance, $X + Y$ aussi et on a $\text{V}\,(X + Y) = \text{V}\,(X) + \text{V}\,(Y)$.

Définition 14 - 7

Si $(X_1, \ldots, X_n)$ sont des variables aléatoires discrètes à valeurs dans E, on appelle n-uplet de variables aléatoires (discrètes) la variable aléatoire discrète à valeurs dans E^n donnée par $(X_1, \ldots, X_n)$.

Si E est un espace vectoriel, X est appelé vecteur aléatoire discret. Lorsque E est de dimension finie, la donnée d'une base de E permet de ramener l'étude d'un vecteur aléatoire discret à celle d'un n-uplet de variables aléatoires discrètes.

Définition 14 - 8

On dit qu'une famille $(X_i)_{i\in I}$ de variables aléatoires discrètes est formée de variables mutuellement indépendantes si, pour toute famille $(A_i)_{i\in I}$ de parties de E, les événements $((X_i \in A_i))_{i\in I}$ sont mutuellement indépendants, i.e. pour tout sous-ensemble fini J de I, on a

$$\mathbf{P}\left(\bigcap_{j\in J}(X_j \in A_j)\right) = \prod_{j\in J}\mathbf{P}(X_j \in A_j)\ .$$

Remarque 14 - 8

Si $(X_1, \ldots, X_n)$ est un n-uplet de variables aléatoires discrètes, ces variables sont mutuellement indépendantes si et seulement si, pour tout n-uplet $(x_1, \ldots, x_n)$ dans $X_1(\Omega) \times \cdots \times X_n(\Omega)$, on a

$$\mathbf{P}(X_1 = x_1, \ldots, X_n = x_n) = \prod_{i=1}^{n}\mathbf{P}(X_i = x_i)\ .$$

Exemple 14 - 3

Si $(X_1, \ldots, X_n)$ est un n-uplet de variables aléatoires mutuellement indépendantes suivant chacune une loi de BERNOULLI de paramètre p, alors $X_1 + \cdots + X_n$ suit une loi binomiale de paramètres n et p.

Proposition 14 - 3

Si $(X_1, \ldots, X_n)$ est un n-uplet de variables aléatoires discrètes à valeurs réelles, chacune admettant une variance, alors leur somme admet une variance et on a

$$\mathrm{V}\left(X_1 + \cdots + X_n\right) = \sum_{k=1}^{n}\mathrm{V}\left(X_k\right) + 2\sum_{1\leq i<j\leq n}\mathrm{Cov}\left(X_i, X_j\right)\ .$$

En particulier si les variables sont deux à deux indépendantes, alors

$$\mathrm{V}\left(X_1 + \cdots + X_n\right) = \mathrm{V}\left(X_1\right) + \cdots + V(X_n)\ .$$

Démonstration. L'existence de la variance résulte du fait que $L^2(\Omega, \mathcal{A}, \mathbf{P})$ est un espace vectoriel, et la formule résulte de la linéarité de l'espérance. $\square$

Exemple 14 - 4

Puisqu'une somme de variables de BERNOULLI de paramètre p suit une loi binomiale de paramètres n et p, si $X \sim \mathcal{B}(n,p)$, alors $\mathrm{V}(X) = n\mathrm{V}(Y)$ où $Y \sim \mathcal{B}(p)$ et donc $\mathrm{V}(X) = np(1-p)$.

Proposition 14 - 4

Lemme des coalitions

Si X_1, X_2, ..., X_n sont des variables aléatoires mutuellement indépendantes, alors pour tout m compris entre 1 et $n-1$, et toutes fonctions f et g, les variables $f(X_1, \ldots, X_m)$ et $g(X_{m+1}, \ldots, X_n)$ sont indépendantes.

Démonstration. On pose $Y = f(X_1, \ldots, X_m)$ et $Z = g(X_{m+1}, \ldots, X_n)$. Pour y dans l'image de f et z dans l'image de g, on pose $A = f^{-1}(y) \cap X_1(\Omega) \times \cdots \times X_m(\Omega)$ et $B = f^{-1}(z) \cap X_{m+1}(\Omega) \times \cdots \times X_n(\Omega)$. Puisqu'on a affaire à des variables discrètes, A

et B sont au plus dénombrables et donc

$$(Y=y)=\sum_{a\in A}((X_1,\dots,X_m)=a) \quad \text{et} \quad (Z=z)=\sum_{b\in B}((X_{m+1},\dots,X_n)=b) \ .$$

Il en résulte

$$(Y=y,Z=z)=\sum_{(a,b)\in A\times B}((X_1,\dots,X_m)=a,(X_{m+1},\dots,X_n)=b) \ .$$

Or, si a est dans A avec $a=(a_1,\dots,a_m)$ et b est dans B avec $b=(b_{m+1},\dots,b_n)$, on a, par indépendance,

$$\begin{aligned}
& \mathbf{P}\left((X_1,\dots,X_m)=a,(X_{m+1},\dots,X_n)=b\right) \\
= \ & \mathbf{P}(X_1=a_1,\dots,X_m=a_m,X_{m+1}=b_{m+1},\dots,X_n=b_n) \\
= \ & \mathbf{P}(X_1=a_1)\times\cdots\times\mathbf{P}(X_m=a_m)\times\mathbf{P}(X_{m+1}=b_{m+1})\times\cdots\times\mathbf{P}(X_n=b_n) \\
= \ & \mathbf{P}(X_1=a_1,\dots,X_m=a_m)\mathbf{P}(X_{m+1}=b_{m+1},\dots,X_n=b_n) \\
= \ & \mathbf{P}\left((X_1,\dots,X_m)=a\right)\mathbf{P}\left((X_{m+1},\dots,X_n)=b\right)
\end{aligned}$$

et donc, par σ-additivité,

$$\begin{aligned}
\mathbf{P}(Y=y,Z=z) &= \sum_{(a,b)\in A\times B}\mathbf{P}\left((X_1,\dots,X_m)=a\right)\mathbf{P}\left((X_{m+1},\dots,X_n)=b\right) \\
&= \left(\sum_{a\in A}\mathbf{P}\left((X_1,\dots,X_m)=a\right)\right)\left(\sum_{b\in B}\mathbf{P}\left((X_{m+1},\dots,X_n)=b\right)\right) \\
&= \mathbf{P}(Y=y)\mathbf{P}(Z=z)
\end{aligned}$$

et donc Y et Z sont indépendants. □

On a déjà évoqué la construction d'une suite de variables aléatoires indépendantes suivant une loi de BERNOULLI de paramètre p grâce à $\{0,1\}^{\mathbf{N}^*}$ et la construction d'une probabilité sur la tribu engendrée par les cylindres (i.e. par les parties déterminées par un nombre fini de tirages) grâce au théorème de CARATHÉODORY. Plus généralement, on admet le théorème suivant.

Théorème 14 - 1

Existence de suites de variables aléatoires
Soit $(\mathcal{L}_k)_{k\in\mathbf{N}}$ une suite de lois discrètes. Alors il existe une suite $(X_k)_{k\in\mathbf{N}}$ de variables aléatoires discrètes, mutuellement indépendantes et telles que, pour tout k dans $\mathbf{N}$, on ait $X_k\sim\mathcal{L}_k$. Autrement dit il existe un espace probabilisé $(\Omega,\mathcal{A},\mathbf{P})$ et X un vecteur aléatoire discret défini sur $(\Omega,\mathcal{A},\mathbf{P})$ avec $X=(X_k)_{k\in\mathbf{N}}$ et, pour tout k dans $\mathbf{N}$, $X_k\sim\mathcal{L}_k$.

En prenant toujours les mêmes lois, on parle de variables aléatoires indépendantes et identiquement distribuées. Par exemple, la modélisation du jeu de *pile ou face* emploie une suite de variables aléatoires indépendantes et identiquement distribuées selon une loi de BERNOULLI de paramètre p.

Exercice

Dans une suite de lancers à *pile ou face* avec une pièce biaisée, quelle est la longueur moyenne de la première séquence homogène ? Et de la seconde ?
Par exemple, si on tire $(PPPPPPPFFFPPF\dots)$ la première séquence est de longueur 7 et la seconde est de longueur 3.

On modélise par une suite $(X_k)_{k\in\mathbf{N}^*}$ de variables aléatoires indépendantes et identiquement distribuées selon une loi de BERNOULLI de paramètre p. On note L_1 et L_2 les deux longueurs (aléatoires) recherchées.

La loi conditionnelle de L_1 sachant $(X_1 = 1)$ est géométrique de paramètre $1-p$ tandis que la loi conditionnelle de L_1 sachant $(X_1 = 0)$ est géométrique de paramètre p. D'après le théorème de sommation par paquets dans le cas positif, on a donc $\mathbf{E}(L_1) = \dfrac{p}{1-p} + \dfrac{1-p}{p}$.

Par contre le même calcul montre que la loi conditionnelle de L_2 sachant $(X_1 = 1)$ est géométrique de paramètre p tandis que la loi conditionnelle de L_2 sachant $(X_1 = 0)$ est géométrique de paramètre $1-p$. Il vient alors $\mathbf{E}(L_2) = \dfrac{p}{p} + \dfrac{1-p}{1-p} = 2$. Par conséquent la seconde séquence homogène est en moyenne toujours plus courte que la première, avec égalité uniquement dans le cas non-biaisé. Elle est même de longueur indépendante de p : le biais éventuel du premier lancer est compensé exactement par le fait qu'on tente d'obtenir la face opposée à celle tirée au premier lancer.

Proposition 14 - 5

Soit X et Y deux variables aléatoires indépendantes à valeurs dans $\mathbf{N}$. Alors le rayon de convergence de la série définissant G_{X+Y} est supérieur au plus petit de ceux correspondant à G_X et G_Y et on a $G_{X+Y} = G_X \cdot G_Y$, quand le membre de droite défini.

Démonstration. On considère le produit de CAUCHY des deux séries entières définissant G_X et G_Y, i.e. $\sum \mathbf{P}(X = n) z^n$ et $\sum \mathbf{P}(Y = n) z^n$. Puisque, pour n et k dans $\mathbf{N}$ avec $k \leq n$, on a par indépendance de X et Y,

$$\mathbf{P}(X = k)\,\mathbf{P}(Y = n-k) = \mathbf{P}(X = k, Y = n-k) = \mathbf{P}(X + Y = n, X = k)$$

et donc

$$\sum_{k=0}^{n} \mathbf{P}(X = k)\,\mathbf{P}(Y = n-k) = \mathbf{P}(X + Y = n)$$

puisque $(X + Y = n) = \displaystyle\sum_{k=0}^{n} (X + Y = n, X = k)$. Il en résulte que la série entière définissant G_{X+Y} est le produit de CAUCHY de celles définissant G_X et G_Y.

L'assertion en découle. □

Remarque 14 - 9

Puisque $G_X(t) = \mathbf{E}\left(t^X\right)$, la propriété s'écrit $\mathbf{E}\left(t^{X+Y}\right) = \mathbf{E}\left(t^X t^Y\right) = \mathbf{E}\left(t^X\right)\mathbf{E}\left(t^Y\right)$, ce qui résulte de l'indépendance de X et Y, et du lemme des coalitions.

On en déduit par récurrence immédiate :

Proposition 14 - 6

Soit $(X_i)_{1\leq i\leq n}$ des variables aléatoires à valeurs dans $\mathbf{N}$, mutuellement indépendantes et Y la somme de ces variables aléatoires, i.e. $y = \displaystyle\sum_{i=1}^{n} X_i$. Alors la série entière définissant G_Y est le produit de CAUCHY de celles définissant G_{X_1}, ..., G_{X_n}, de sorte que, lorsque les termes de droite sont définis, on a

$$G_Y(t) = G_{X_1+\cdots+X_n}(t) = G_{X_1}(t)\cdots G_{X_n}(t) = \prod_{i=1}^{n} G_{X_i}(t)\,.$$

2 Phénomènes limites – Loi faible des grands nombres

Soit $(X_n)_{n\in\mathbf{N}^*}$ une suite de variables aléatoires indépendantes et identiquement distribuées de loi $\mathcal{B}(p)$. On note $S_n = \sum_{k=1}^{n} X_k$ et $M_n = \frac{1}{n}S_n$, les sommes et moyennes. Dans ce cadre, $S_n \sim \mathcal{B}(n,p)$ et donc $\mathbf{E}(M_n) = \frac{1}{n}\mathbf{E}(S_n) = p$ et $\mathrm{V}(M_n) = \frac{1}{n^2}\mathrm{V}(S_n) = \frac{p(1-p)}{n}$. Autrement dit la suite $(M_n - p)_{n\in\mathbf{N}^*}$ est une suite de variables centrées dont la variance tend 0. Comme on l'a déjà remarqué, l'inégalité de Bienaymé-Tchebychev entraîne la convergence de $(M_n - p)$ vers 0 en probabilité, et donc, pour tout ε dans $\mathbf{R}_+^*$,

$$\lim \mathbf{P}\left(|M_n - p| > \varepsilon\right) = 0\ .$$

Ce résultat a été obtenu par Jacob Bernoulli. Plus généralement

Théorème 14 - 2

Loi faible des grands nombres

Soit $(X_n)_{n\in\mathbf{N}^*}$ une suite de variables aléatoires deux à deux indépendantes, de même loi et admettant un moment d'ordre 2, alors, en posant $S_n = \sum_{k=1}^{n} X_k$ et $\mu = \mathbf{E}(X_1)$, on a

$$\mathbf{P}\left(\left|\frac{S_n}{n} - \mu\right| \geq \varepsilon\right) \underset{n\to+\infty}{\longrightarrow} 0\ .$$

Démonstration. Par linéarité de l'espérance, pour n dans $\mathbf{N}^*$, on a $\mathbf{E}\left(\frac{S_n}{n}\right) = \mu$.

Par indépendance des (X_n), on a, pour n dans $\mathbf{N}^*$, $\mathrm{V}(S_n) = n\mathrm{V}(X_1)$ et donc $\mathrm{V}\left(\frac{S_n}{n}\right) = \frac{\mathrm{V}(X_1)}{n}$. L'inégalité de Bienaymé-Tchebychev permet donc de conclure, pour ε dans $\mathbf{R}_+^*$,

$$\mathbf{P}\left(\left|\frac{S_n}{n} - \mu\right| \geq \varepsilon\right) \leq \frac{\mathrm{V}(X_1)}{n\varepsilon^2}$$

et le résultat en découle. □

Remarque 14 - 10

On n'utilise pas, en fait, que les variables sont identiquement distribuées, simplement qu'elles ont même espérance et même variance. C'est donc uniquement une restriction du programme.

Pour aller plus loin

On peut démontrer que le résultat précédent est encore vrai sans hypothèse sur l'existence de moments d'ordre 2, i.e. si $(X_n)_{n\in\mathbf{N}^*}$ une suite de variables aléatoires deux à deux indépendantes, de même loi, alors leur moyenne converge en probabilité vers la variable presque sûrement égale à leur espérance commune μ.

Dans ce même cadre, Kolmogorov a démontré un résultat encore plus précis : la moyenne converge presque sûrement vers μ. Il s'agit de la **loi forte des grands nombres**.

3 Compléments

Supposons qu'on ait $\Omega = \mathbf{R}$ et que les ensembles $(X = x)$, pour x dans $X(\Omega)$ soient des réunions finies d'intervalles de $\mathbf{R}$ avec $\mathbf{P}(X = x) = \int_{(X=x)} f$ pour une certaine fonction f positive d'intégrale 1. On peut alors écrire

$$X = \sum_{x \in X(\Omega)} x \mathbb{1}_{(X=x)}$$

et donc X est une fonction en escalier. Son espérance est alors son intégrale

$$\mathbf{E}(X) = \int_{\mathbf{R}} X(t) f(t)\, \mathrm{d}t$$

et cette intégrale n'est pas seulement prise comme la limite des intégrales sur des segments quand les bornes de ces segments tendent vers l'infini : elle n'est définie que si X est une fonction intégrable (i.e. $|X|$ l'est). Ces précautions sont nécessaires pour avoir affaire à des espaces vectoriels normés.

On dit d'ailleurs qu'une variable aléatoire qui admet un moment d'ordre 1 est intégrable.

La convergence en loi, i.e. lorsque la loi de X_n converge vers celle de X est une notion très faible et n'implique pas de relation forte entre les (X_n) et X, seulement leurs lois. Il est équivalent de demander, pour toute fonction continue et bornée de $\mathbf{R}$ dans $\mathbf{R}$, $\lim \mathbf{E}(f(X_n)) = \mathbf{E}(f(X))$.

♠

Soit $(X_n)_{n\in\mathbf{N}}$ une suite de variables aléatoires à valeurs dans $\mathbf{N}$, et X une telle variable aléatoire. Alors X_n converge en loi vers X si et seulement si l'une des deux propriétés équivalentes suivantes est vérifiée :

1. Pour tout k dans $\mathbf{N}$, $\lim \mathbf{P}(X_n = k) = \mathbf{P}(X = k)$;
2. La suite de fonctions $(G_{X_n})_{n\in\mathbf{N}}$ converge simplement vers G_X sur $[0;1[$.

Danger

La convergence en loi ne permet donc que d'évaluer des probabilités et non de comparer des événements reliés aux variables aléatoires en jeu. Néanmoins on peut construire des variables aléatoires (Y_n) et Y, avec $Y_n \sim X_n$ et $Y \sim X$ et $Y_n \xrightarrow{ps} Y$.

Pour aller plus loin

La convergence en loi est en fait équivalente à la convergence simple des fonctions caractéristiques définies par $\varphi_X = \mathbf{E}\left(e^{itX}\right)$ ou, plus généralement, à la convergence de $\mathbf{E}(f(X_n))$ vers $\mathbf{E}(f(X))$ pour toute fonction f continue et bornée.

La convergence en probabilité est équivalente à la convergence de $\mathbf{E}(f(|X_n - X|))$ vers 0 pour une fonction f continue, bornée, positive, nulle en 0, croissante et strictement croissante dans un voisinage de 0. Par exemple

$$\lim \mathbf{E}\left(\frac{|X_n - X|}{1 + |X_n - X|}\right) = 0 \quad \text{ou} \quad \lim \mathbf{E}(\min(|X_n - X|, 1)) = 0\,.$$

Exercices

Famille de variables aléatoires

14 - 1 ★ Variables de Bernoulli

Montrer que deux variables de Bernoulli sont indépendantes si et seulement si leur covariance est nulle.

14 - 2 ★ Jet de dés

On jette deux dés équilibrés. Trouver la loi conjointe de X et Y dans les cas suivants :

a. X est la plus grande des deux valeurs obtenues et Y en est la somme.

b. X est la valeur obtenue avec le premier dé et Y est la plus grande des deux valeurs obtenues.

c. X et Y sont la plus petite et la plus grande des deux valeurs obtenues.

14 - 3 Ⓢ ★ Temps d'arrêt

Soit $(X_k)_{k\in\mathbf{N}^*}$ une famille de variables aléatoires indépendantes et identiquement distribuées, à valeurs dans $[\![0;n]\!]$, avec $\mathbf{P}(X_1=0) < 1$. On note $N = \inf\{k\in\mathbf{N}^* \mid X_k \neq 0\}$ et $X = X_N$.

a. Montrer que N et X sont des variables aléatoires.

b. Déterminer la loi de N.

c. Déterminer celle de X.

d. Montrer $\mathbf{P}(N=k, X=j) = \mathbf{P}(N=n)\,\mathbf{P}(X=j)$.

e. L'indépendance de N par rapport à X paraît-elle intuitive ?

f. L'indépendance de X par rapport à N paraît-elle intuitive ?

14 - 4 Ⓢ M 2015 ★★ Sommes de variables de Bernoulli

Soit (X_n) une suite de variables aléatoires indépendantes suivant une loi de Bernoulli de paramètre p dans $]0;1[$.

a. On pose, pour k entier, $Y_k = X_k + X_{k+1}$. Donner la loi de Y_k, son espérance et sa variance, ainsi que la covariance de Y_i et Y_j pour $i\neq j$.

b. On pose, pour n entier, $T_n = \dfrac{1}{n}\displaystyle\sum_{k=1}^{n} Y_k$. Calculer son espérance et sa variance.

14 - 5 Ⓢ ★★ Théorème de Kantorovitch

Soit X et Y des variables aléatoires réelles prenant respectivement m et n valeurs. Montrer que X et Y sont indépendantes si et seulement si $\operatorname{Cov}\left(X^k, Y^\ell\right) = 0$ pour k et ℓ entiers avec $k<m$ et $\ell<n$.

14 - 6 Ⓢ ★★ Convolution ♥

Soit X et Y deux variables aléatoires à valeurs dans $\mathbf{Z}$ et indépendantes.

a. Montrer que $X+Y$ est une variable aléatoire et que sa loi est donnée par

$$\mathbf{P}(X+Y=n) = \sum_{k\in\mathbf{Z}} \mathbf{P}(X=k)\mathbf{P}(Y=n-k)\ .$$

On note $\mathcal{L}_{X+Y} = \mathcal{L}_X \star \mathcal{L}_Y$.

b. Montrer $\mathcal{P}(\lambda)\star\mathcal{P}(\mu) = \mathcal{P}(\lambda+\mu)$.

c. Motrer $\mathcal{B}(n,p)\star\mathcal{B}(m,p) = \mathcal{B}(n+m,p)$.

14 - 7 Ⓢ ★★ Quotient de variables géométriques

Soit X et Y deux variables aléatoires indépendantes avec $X\sim Y\sim\mathcal{G}(p)$ et $Z = X/Y$.

a. Montrer que Z est une variable aléatoire et en déterminer la loi.

b. En utilisant la formule de transfert en déterminer l'espérance.

14 - 8 Ⓢ ★★ Maximum de variables géométriques

Soit X et Y deux variables aléatoires indépendantes avec $X\sim Y\sim\mathcal{G}(p)$ et $Z = \max(X,Y)$.

a. Montrer que Z est une variable aléatoire et en déterminer la loi.

b. En déterminer l'espérance.

14 - 9 Ⓢ ★★★ Somme aléatoire de variables de Bernoulli

On effectue une succession de lancers à pile ou face avec une probabilité p d'obtenir pile. Soit N le nombre de tirages nécessaires pour obtenir un pile. On tire à nouveau N fois la même pièce. Montrer qu'en moyenne on obtient alors un seul pile. Autrement dit si $(X_k)_{k\in\mathbf{N}^*}$ est une suite de variables aléatoires indépendantes avec $X_k\sim\mathcal{B}(p)$ et si $N = \inf\{n\in\mathbf{N}^* \mid X_1+\cdots+X_n = 1\}$, alors N est une variable aléatoire et on a $\mathbf{E}(X_{N+1}+\cdots+X_{2N}) = 1$.

14 - 10 Ⓢ ★★★ Tri rapide

Soit $(x_k)_{1\leq k\leq n}$ une famille de réels tous distincts que l'on souhaite ordonner. On initialise l'algorithme en fabriquant une liste constituée d'une unique liste de taille n avec ces réels. À chaque étape de l'algorithme on effectue les étapes suivantes :

1. Si la liste ne contient que des listes de taille 1, on s'arrête ;
2. Sinon on choisit au hasard, selon une loi uniforme, l'une des listes de taille supérieure à 1, notée L ;

3. Dans cette liste on choisit un terme x au hasard, selon une loi uniforme et on remplace la liste L, par trois listes : les termes de L inférieurs à x, x, et les termes strictement supérieurs à x.

On note X le nombre de comparaisons nécessaires avant que l'algorithme ne se termine et on s'intéresse à $\mathbf{E}(X)$.

a. Montrer que la probabilité que x_j et x_k soient comparés durant l'algorithme est égale à $\dfrac{2}{N}$ où N est le nombre de réels de la famille compris entre x_j et x_k, i.e. $N = N(i,j) = \operatorname{Card}\{i \in [\![1;n]\!] \mid (x_i - x_j)(x_k - x_i) \geq 0\}$.

b. Montrer $\mathbf{E}(X) = \displaystyle\sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \frac{2}{j-i+1}$. (On pourra ordonner les réels et écrire X comme une somme d'indicatrices.)

c. En déduire qu'on a $\mathbf{E}(X) \sim 2n\ln(n)$.

Lois conditionnelles

14 - 11 Ⓢ ★ Somme de variables de Poisson

Soit X et Y deux variables aléatoires indépendantes suivant des lois de Poisson de paramètres respectifs λ et μ. Calculer la loi de X sachant $X + Y = n$.

14 - 12 Ⓢ ★ Probabilité des causes ♥

Soit $(B_n)_{n\in I}$ un système complet d'événements tous non-négligeables, i.e. une partition de Ω avec I au plus dénombrable et vérifiant $\mathbf{P}(B_n) > 0$ pour n dans I, et X une variable aléatoire discrète réelle.

a. Soit A un événement observable. Montrer

$$\inf_{n\in I} \mathbf{P}(A \mid B_n) \leq \mathbf{P}(A) \leq \sup_{n\in I} \mathbf{P}(A \mid B_n) \ .$$

b. On suppose que la réalisation de l'événement A double la probabilité de réalisation d'un événement B_n, i.e. $\dfrac{\mathbf{P}(B_n \mid A)}{\mathbf{P}(B_n)} = 2$. Que dire de la probabilité de réalisation de A selon que l'on sait ou pas que B_n est réalisé ? Qu'en déduire en terme de causalité ?

14 - 13 Ⓢ ★★ Probabilités conditionnelles ♥♥

Soit A un événement observable non-négligeable et X une variable aléatoire.

a. Montrer que $\left\{\emptyset, A, \overline{A}, \Omega\right\}$ est une tribu $\sigma(A)$ et donner la restriction $\mathbf{P}_A$ de $\mathbf{P}$ à cette tribu. (On veillera à ne pas confondre $\mathbf{P}_A$ avec la probabilité conditionnelle selon A.)

b. Montrer que l'espace vectoriel réel E_A engendré par $\mathbb{1}_A$ et $\mathbb{1}_{\overline{A}}$ (en tant que sous-espace de celui des fonctions de Ω dans $\mathbf{R}$) est égal à l'algèbre des variables aléatoires sur $(\Omega, \sigma(A), \mathbf{P}_A)$.

c. On considère E l'espace vectoriel de dimension au plus 3 engendré par X et E_A et on le munit du produit scalaire donné par $\langle u \mid v\rangle = \mathbf{E}(uv)$. On note Y la variable aléatoire obtenue par projection orthogonale de X sur E_A. Montrer

$$Y = \mathbf{E}(X \mid A)\,\mathbb{1}_A + \mathbf{E}\left(X \;\middle|\; \overline{A}\right)\mathbb{1}_{\overline{A}} \,.$$

d. En particulier, lorsque $X = \mathbb{1}_B$, donner une interprétation de $\mathbf{P}(B \mid A)$.

14 - 14 Ⓢ ★★ Somme aléatoire de variables de Poisson

On admet que le nombre de client(e)s dans un bureau de poste durant une journée est une variable aléatoire poissonnienne de paramètre λ et que la probabilité pour qu'une personne pénétrant dans le bureau de poste vienne pour une opération postale est p, tandis que cette probabilité est $1-p$ pour les opérations bancaires. Montrer que le nombre X de client(e)s durant une journée qui viennent pour une opération postale et le nombre Y de ces mêmes client(e)s venant pour une opération bancaires sont des variables aléatoires poissonniennes de paramètres λp et $\lambda(1-p)$ respectivement, et que ces deux variables aléatoires sont de plus indépendantes.

14 - 15 Ⓢ ★★★ Espérance conditionnelle ♥

Soit B un évenement non négligeable et X une variable aléatoire discrète réelle. On note $\mathbf{E}(X \mid B) = \displaystyle\sum_{x\in X(\Omega)} x\mathbf{P}(X = x \mid B)$, i.e. l'espérance de X vu comme variable aléatoire sur $(\Omega, \mathcal{A}, \mathbf{P}_B)$. Soit $(B_n)_{n\in I}$ un système complet d'événements tous non-négligeables, i.e. une partition de Ω avec I au plus dénombrable et vérifiant $\mathbf{P}(B_n) > 0$ pour n dans I.

a. Montrer le théorème de l'espérance mathématique totale

$$\mathbf{E}(X) = \sum_{n\in I} \mathbf{P}(B_n)\mathbf{E}(X \mid B_n) \ .$$

b. Soit $(X_i)_{i\in\mathbf{N}^*}$ une suite de variables aléatoires indépendantes et identiquement distribuées selon une loi $\mathcal{B}(p)$. On note T_1 le temps d'attente avant le premier succès, i.e. $T_1 = \min\{i \in \mathbf{N}^* \mid X_i = 1\}$. On fixe n dans $\mathbf{N}^*$ et on pose $T_1^{(n)} = \min(T_1, n+1)$ et $S_n = \displaystyle\sum_{i=1}^{n} X_i$. Montrer que S_n, T_1 et $T_1^{(n)}$ sont des variables aléatoires et qu'on a $\mathbf{E}(T_1 \mid S_n = k) = \dfrac{n+1}{k+1}$.

<u>Indication</u> : Avec des notations évidentes, on pourra montrer que $(T_1, \ldots, T_{k+1})$ ont même loi.

c. Montrer $\mathbf{E}\left(T_1^{(n)}\right) = \displaystyle\sum_{k=0}^{n} \mathbf{P}\left(T_1^{(n)} > k\right)$ et en déduire

$$\mathbf{E}\left(\frac{1}{S_n + 1}\right) = \frac{1-(1-p)^{n+1}}{(n+1)p} \ .$$

14 - 16 X 2015 ★★★ Espérance conditionnelle

Soit X_1 et X_2 deux variables aléatoires à valeurs dans $\mathbf{Z}$. On suppose que X_2 admet une espérance.

a. Montrer qu'il existe une fonction h de $\mathbf{Z}$ dans $\mathbf{R}$, telle que, en notant $Y_1 = h(X_1)$, Y_1 admet une espérance et, pour toute application bornée f de $\mathbf{Z}$ dans $\mathbf{R}$,
$$\mathbf{E}(X_2 f(X_1)) = \mathbf{E}(Y_1 f(X_1)) \ .$$

b. Montrer l'unicité au sens suivant : si k est une autre fonction vérifiant la propriété précédente, alors $h(X_1) = k(X_1)$ presque sûrement.

c. On note $Y_1 = \mathbf{E}(X_1 \mid X_2)$, c'est l'espérance conditionnelle de X_1 par rapport à X_2. Soit g une application bornée de $\mathbf{Z}$ dans $\mathbf{R}$. Montrer
$$\mathbf{E}(X_2 g(X_1) \mid X_1) = g(X_1)\mathbf{E}(X_2 \mid X_1) \ .$$

Fonctions génératrices

14 - 17 Ⓢ M 2018 ★★ Sommes de variables géométriques

a. Soit $(X_n)_{n\in\mathbf{N}^*}$ une suite de variables aléatoires indépendantes suivant une loi géométrique de paramètre p dans $]0;1[$. Déterminer la loi de la somme d'un nombre fini d'entre elles.

b. Une bactérie a une probabilité p de se faire toucher par un laser, et il faut n coups de laser pour la tuer. Calculer la loi de X, temps de vie de la bactérie, et son espérance sachant qu'il y a une minute entre chaque passage de laser.

14 - 18 Ⓢ ★★ Piper des dés

a. Montrer que le polynôme $\displaystyle\sum_{k=0}^{10} X^k$ ne peut pas s'écrire comme un produit de deux polynômes à coefficients réels, tous deux de degré inférieur à 5.

b. En utilisant les fonctions génératrices, en déduire qu'il est impossible de piper deux dés (indépendamment) de telle sorte que la somme des valeurs de ces dés lors d'un tirage (indépendant) suive une loi uniforme sur $[\![2;12]\!]$.

14 - 19 Ⓢ ★★★ Somme aléatoire

Soit $(X_n)_{n\in\mathbf{N}^*}$ des variables aléatoires indépendantes et identiquement distribuées, à valeurs entières, et N une variable aléatoire indépendante des précédentes, également à valeurs entières. On pose $S = \displaystyle\sum_{n=1}^{N} X_n$, i.e. pour tout ω dans Ω,
$$S(\omega) = \sum_{n=1}^{N(\omega)} X_n(\omega) \ .$$
On note G la fonction génératrice commune des (X_n). Montrer, pour $t \in [-1;1]$, $G_S(t) = G_N(G(t))$.

14 - 20 Ⓢ ★★★ Fonction génératrice des moments

Soit X une variable aléatoire à valeurs dans $\mathbf{N}$. On pose $H(t) = G_X(e^t) = \mathbf{E}\left(e^{tX}\right)$ et $I(t) = e^{-t\mathbf{E}(X)}H(t) = \mathbf{E}\left(e^{t(X-\mathbf{E}(X))}\right)$.

a. Montrer, sous réserve d'existence, que H est la somme de la série entière $\displaystyle\sum \frac{\mathbf{E}\left(X^k\right)}{k!}t^k$ et que I est la somme de la série entière $\displaystyle\sum \frac{\mathbf{E}\left((X-\mathbf{E}(X))^k\right)}{k!}t^k$.

b. Montrer que H et I sont bien définies dans un voisinage de 0 si et seulement si le rayon de convergence de G_X est strictement supérieur à 1, et que cette condition équivaut à l'existence de moments de tous ordres pour X et que le rayon de convergence de la série entière $\displaystyle\sum \frac{\mathbf{E}\left(X^k\right)}{k!}t^k$ est non nul.

14 - 21 Ⓢ ★★★ Processus de Galton-Watson

Soit $(X_i^{(n)})_{(n,i)\in\mathbf{N}\times\mathbf{N}^*}$ des variables aléatoires indépendantes et identiquement distribuées, à valeurs entières.

Avec les notations de l'exercice 14 - 19, on pose $Z_0 = 1$ et, pour $n \in \mathbf{N}$,
$$Z_{n+1} = \sum_{i=1}^{Z_n} X_i^{(n)} \ .$$

a. Montrer que $(Z_n)_{n\in\mathbf{N}}$ est une suite de variables aléatoires à valeurs dans $\mathbf{N}$.

b. En utilisant le résultat de l'exercice 14 - 19 et en notant G la fonction génératrice de Z_1, montrer $G_{Z_n} = G^n$, où l'exposant désigne la composition des fonctions : $G^0 = \mathrm{Id}$, $G^2 = G \circ G$ etc.

c. On note $x_n = \mathbf{P}(Z_n = 0)$. Démontrer que la suite $(x_n)_{n\in\mathbf{N}}$ converge vers le plus petit point fixe de G sur $[0;1]$.

d. On suppose que $X_i^{(n)}$ représente la descendance de l'individu numéroté i et appartenant à la n^{e} génération.

Par exemple le nombre de neutrons libérés lors de la collision d'un neutron avec un noyau atomique, ou le nombre de personnes infectées lors de la propagation d'une épidémie.

Interpréter la question précédente à l'aide de l'événement : la propagation s'arrête.

e. Caractériser le cas où, presque sûrement, la propagation continue indéfiniment.

14 - 22 ★★★ Processus de Galton-Watson

On reprend l'exercice 14 - 21 et on suppose de plus que les $X_i^{(n)}$ admettent des moments d'ordre 2. On note μ et σ leurs moyenne et écart-type.

a. En utilisant la croissance et la convexité de G, montrer que l'arrêt de la propagation est presque sûr si et seulement si $\mu < 1$ ou bien $\mu = 1$ et $\sigma > 0$.

b. Que dire si $\mu = 1$ et $\sigma = 0$?

c. Que dire si $\mu > 1$?

d. Déterminer $\mathbf{E}(Z_n)$ pour tout entier n. Interpréter.

e. Déterminer $\mathrm{V}(Z_n)$ pour tout entier n.

Pile ou Face

14 - 23 Ⓢ M 2017 ★★ Ruine du joueur ♥

On considère un entier naturel non-nul N et n un entier naturel inférieur à N, ainsi qu'un jeu à deux joueurs, A et B. Le joueur A possède initialement n billes et le joueur B en possède $N - n$. Une partie se joue en manches successives. Si un joueur perd une manche, il doit donner une bille à l'autre. La partie se termine quand un joueur possède les N billes, ce joueur étant alors déclaré gagnant. Le joueur A remporte chaque manche avec une probabilité p dans $]0;1[$. Déterminer la probabilité que A gagne.

14 - 24 Ⓢ ★★ Théorème de Weierstrass ♥

Soit $(X_n)_{n\in\mathbf{N}^*}$ une suite de variables aléatoires indépendantes et identiquement distribuées selon une loi $\mathcal{B}(p)$ et f une fonction continue sur $[0;1]$. On pose $M_n = \dfrac{X_1 + \cdots + X_n}{n}$ et $Y_n = f(M_n)$.

a. Soit ε dans $\mathbf{R}_+^*$. On dispose de δ un module d'uniforme continuité pour f sur le segment $[0;1]$ associé à ε. En écrivant

$$|Y_n - f(p)| = |Y_n - f(p)|\,\mathbb{1}_{|M_n - p| > \delta} + |Y_n - f(p)|\,\mathbb{1}_{|M_n - p| \leq \delta}\,,$$

démontrer $\mathbf{E}(|Y_n - f(p)|) \leq \dfrac{\|f\|_\infty}{2n\delta^2} + \varepsilon$.

b. En déduire $\lim \mathbf{E}(|Y_n - f(p)|) = 0$.

c. Plus précisément montrer que $\mathbf{E}(Y_n)$ est une expression polynomiale en p et que la suite de polynômes associée converge uniformément vers f sur $[0;1]$.

d. En déduire que toute fonction continue sur un segment de $\mathbf{R}$ est limite uniforme, sur ce segment, d'une suite de fonctions polynomiales.

14 - 25 Ⓢ ★★ Événements régénératifs

On fixe une suite finie $(x_k)_{1\leq k\leq r}$ à valeurs dans $\{0,1\}$. Soit $(X_n)_{n\in\mathbf{N}^*}$ une suite de variables aléatoires indépendantes et identiquement distribuées selon une loi $\mathcal{B}(p)$. On note $A_0 = \Omega$, $B_0 = \emptyset$, $A_n = B_n = \emptyset$ si $1 \leq n < r$ et enfin, pour $n \geq r$, $A_n = \displaystyle\bigcap_{k=1}^{r}(X_{n-r+k} = x_k)$ et $B_n = A_n \cap \overline{\left(\displaystyle\bigcup_{k=1}^{n-1} A_k\right)}$.

a. Montrer que les événements (A_n) et (B_n) sont observables. On pose $a_n = \mathbf{P}(A_n)$ et $b_n = \mathbf{P}(B_n)$.

b. Montrer que les séries entières $\sum a_n z^n$ et $\sum b_n z^n$ ont un rayon de convergence supérieur ou égal à 1. On note $A(z)$ et $B(z)$ leurs sommes respectives.

c. On dit que l'événement « le tirage donne $(x_k)_{1\leq k\leq r}$ » est régénératif si, pour n et k dans $\mathbf{N}$, on a $\mathbf{P}(A_{n+k} \mid A_n) = \mathbf{P}(A_k)$. Étudier si on a affaire à un événement régénératif lorsque la suite $(x_k)_{1\leq k\leq r}$ est, respectivement, la suite $(1,1,0)$, $(0,1,1)$ ou $(1,1,1)$.

d. À quelle condition sur la suite $(x_k)_{1\leq k\leq r}$ a-t-on affaire à un événement régénératif ?

e. Montrer que si c'est le cas on a $A = 1 + AB$ quand les deux sommes A et B sont définies.

f. Calculer A et B dans les cas d'événements régénératifs étudiés en **c.**, en supposant $p = \dfrac{1}{2}$.

14 - 26 Ⓢ ★★ Temps d'attente

On reprend la situation de l'exercice 14 - 25, avec $p = \dfrac{1}{2}$. On définit T par $(T = n) = B_n$ et $(T = +\infty) = \Omega \setminus \bigcup_{n=0}^{+\infty} B_n$.

a. Montrer que T définit une variable aléatoire presque sûrement finie. Interpréter T.

b. Montrer $\mathbf{E}(T) = B'(1)$, cette égalité étant valable dans $\mathbf{R}_+ \cup \{+\infty\}$.

c. Calculer cette espérance dans les cas $(1,1,0)$ et $(0,1,1)$.

d. Vaut-il mieux parier que $(1,1,0)$ sortira avant $(0,1,1)$ plutôt que le contraire ?

14 - 27 ★★★ Cas non régénératif

On reprend la situation de l'exercice 14 - 26 dans le cas $(1,1,1)$.

a. Montrer que l'événement « le tirage donne 1 consécutivement un nombre de fois égal à un multiple de 3 » est régénératif.

b. On note a_n la probabilité que cet événement ait lieu au n^{e} tirage (on convient $a_0 = 1$). Montrer, pour $n \geq 3$, $a_n + \dfrac{1}{2}a_{n-1} + \dfrac{1}{4}a_{n-2} = \dfrac{1}{8}$.

c. Montrer que, à l'intérieur de son domaine de convergence, la somme de la série entière $\sum a_n z^n$ est donnée par $A(z) = \dfrac{16 - 16z + z^4}{2(1-z)(8-z^3)}$.

d. On définit B par $A = 1 + AB$. Interpréter B.

e. Calculer $B(1)$. Interpréter.

f. Calcul $B'(1)$. Interpréter.

14 - 28 ★★★ Retour à l'équilibre

On reprend la situation de l'exercice 14 - 26.

a. Montrer que l'événement « le tirage donne autant de 1 que de 0 » est régénératif.

b. On note a_n la probabilité que cet événement ait lieu au n^e tirage (on convient $a_0 = 1$). Montrer que, à l'intérieur de son domaine de convergence, la somme de la série entière $\sum a_n z^n$ est donnée par $A(z) = (1 - z^2)^{-1/2}$ (pour z réel).

c. En déduire que le retour à l'équilibre est presque sûr, i.e. que l'événement « il existe un entier n dans $\mathbf{N}^*$ tel qu'il y a autant de 1 que de 0 dans la suite $(X_k)_{1\leq k\leq n}$ » est certain. On pourra introduire B donné par $A = 1 + AB$.

d. Montrer que, presque sûrement, il y a un nombre infini de retours à l'équilibre. On pourra considérer l'événement « il y a $n+1$ retours à l'équilibre » conditionnellement à l'événement « il y a n retours à l'équilibre ».

e. On se place dans le cas plus général avec p quelconque.

- **i.** Montrer qu'alors $A(z) = (1 - 4pqz^2)^{-1/2}$.
- **ii.** En déduire que la probabilité qu'il y ait un retour à l'équilibre est $1 - |p - q|$.
- **iii.** En déduire que, si $p \neq q$, alors il n'y a presque sûrement qu'un nombre fini de retours à l'équilibre.

14 - 29 Ⓢ ★★ Une idée de Cantelli ♥

Soit $(X_n)_{n\in\mathbf{N}^*}$ une suite de variables aléatoires indépendantes et identiquement distribuées selon une loi $\mathcal{B}(p)$. On note $Y_n = M_n - p = \dfrac{X_1 + \cdots + X_n}{n} - p$. Montrer qu'on a

$$\mathbf{E}\left(Y_n^4\right) = \frac{1}{n^3}p(1-p)(1-3p+3p^2) + 3\frac{n-1}{n^3}p^2(1-p)^2$$

et en déduire d'une part que la série $\sum \mathbf{E}\left(Y_n^4\right)$ est convergente et d'autre part que, pour tout ε dans $\mathbf{R}_+^*$, l'événement « $|Y_n| > \varepsilon$ infiniment souvent » est négligeable.

14 - 30 Ⓢ ★★ Loi forte des grands nombres ♥

En utilisant les notations et le résultat de l'exercice 14 - 29, démontrer $Y_n \xrightarrow{ps} 0$, i.e. $M_n \xrightarrow{ps} p$. On pourra écrire le complémentaire de l'événement $Y_n \to 0$ comme une réunion dénombrable d'événements négligeables.

Inégalités

14 - 31 Ⓢ ★ Inégalité de Jensen ♥

Soit X une variable aléatoire réelle discrète admettant un moment d'ordre 1 et f une fonction convexe sur un ouvert contenant $X(\Omega)$. En utilisant que f est au-dessus de sa (ou ses) tangente(s) en $\mathbf{E}(X)$, montrer

$$\mathbf{E}(f(X)) \geq f(\mathbf{E}(X)) \ .$$

14 - 32 Ⓢ ★ Inégalité unilatérale de Tchebychev ♥

Soit X une variable aléatoire discrète centrée de variance σ^2 finie. Montrer que pour tous réels strictement positifs a et b on a $\mathbf{P}(X \geq a) \leq \mathbf{P}\left((X+b)^2 \geq (a+b)^2\right)$ et en déduire, en utilisant l'inégalité de Markov, $\mathbf{P}(X \geq a) \leq \dfrac{\sigma^2}{\sigma^2 + a^2}$. Comparer avec l'inégalité de Bienaymé-Tchebychev.

14 - 33 Ⓢ ★ Inégalité de Chernoff ♥

Soit X une variable aléatoire discrète réelle, a un réel et t un réel strictement positif. En utilisant l'inégalité de Markov, montrer $\mathbf{P}(X \geq a) \leq e^{-ta}\mathbf{E}\left(e^{tX}\right)$.

14 - 34 Ⓢ ★★ Application - Inégalité unilatérale

On forme au hasard cent paires à partir d'un groupe initialement formé de cent femmes et cent hommes. Montrer, en utilisant l'inégalité unilatérale de Tchebychev (exercice 14 - 32), que la probabilité pour qu'il y ait moins de trente paires mixtes est inférieure à 6%.

14 - 35 Ⓢ ★★ Application – Inégalité de Chernoff

Soit $(X_n)_{n\in\mathbf{N}^*}$ une suite de variables aléatoires indépendantes et identiquement distribuées prenant les valeurs -1 et 1 avec probabilité $\dfrac{1}{2}$. On pose $S_n = X_1 + \cdots + X_n$. Montrer $\mathbf{E}\left(e^{tS_n}\right) = \mathbf{E}\left(e^{tX_1}\right)^n \leq e^{nt^2/2}$. En utilisant l'inégalité de Chernov (exercice 14 - 33), en déduire $\mathbf{P}(S_n \geq a) \leq e^{-a^2/2n}$ (pour $a > 0$). Donner en particulier une majoration de $\mathbf{P}(S_{10} \geq 6)$ et comparer à la valeur exacte.

14 - 36 Ⓢ ★★★ Collection

On collectionne des cartes d'un jeu. Il y a N sortes de cartes différentes. On suppose que toutes les cartes ont la même chance d'être obtenues, i.e. on se donne $(X_k)_{1\leq k\leq n}$ des variables aléatoires indépendantes et identiquement distribuées de loi $\mathcal{U}(\llbracket 1;N\rrbracket)$, et on s'intéresse à la quantité nécessaire T pour obtenir la collection complète, i.e.

$$T = \inf\left\{n \in \mathbf{N}^* \mid \operatorname{Card}\left\{X_k \mid 1 \leq k \leq n\right\} = N\right\} \ .$$

a. Montrer que T est une variable aléatoire définie presque sûrement.

b. On fixe n et on introduit, pour $1 \leq j \leq N$, l'événement A_j défini par $j \notin \{X_k \mid 1 \leq k \leq n\}$ et, pour $J \subset \llbracket 1;N\rrbracket$, $A_J = \cap_{j\in J} A_j$.

- **i.** Calculer $\mathbf{P}(A_j)$.
- **ii.** Calculer $\mathbf{P}(A_J)$.
- **iii.** ♠ En déduire $\mathbf{P}(T > n)$ en utilisant la formule de Poincaré.

c. Estimer n tel que $\mathbf{P}(T > n)$ soit inférieur à 5% en supposant $N = 20$.

14 - 37 ★★★ Collection (bis)

On reprend l'exercice précédent et on s'intéresse au nombre de cartes différentes obtenues, i.e. $D_n = \operatorname{Card}\{X_k \mid 1 \leq k \leq n\}$.

a. On fixe une partie J de $[\![1;N]\!]$ et on note respectivement A_J et B_J les événements $\{X_k \mid 1 \leq k \leq n\} \subset J$ et $J \subset \{X_k \mid 1 \leq k \leq n\}$.

i. Calculer $\mathbf{P}(A_J)$.

ii. Calculer $\mathbf{P}(B_J \mid A_J)$.

b. En déduire la loi de D_n.

c. Retrouver le résultat de l'exercice précédent.

Convergence

14 - 38 Ⓢ ★ Limite simple

Soit X une variable aléatoire discrète à valeurs dans $\mathbf{R}_+$. Montrer qu'il existe une suite $(X_n)_{n\in\mathbf{N}^*}$ de variables aléatoires prenant chacune un nombre fini de valeurs et vérifiant $\lim \uparrow X_n = X$, i.e. pour tout ω dans Ω, la suite $(X_n(\omega))$ converge en croissant vers $X(\omega)$.

14 - 39 Ⓢ ★ Convergence dominée discrète ♥♥

Soit $\left(\sum a_n^{(k)}\right)_{k\in\mathbf{N}}$ une suite de séries à valeurs réelles. On suppose que, pour tout n dans $\mathbf{N}$, $\lim_k a_n^{(k)} = a_n$ et qu'il existe une série $\sum b_n$ à termes positifs et convergente telle que, pour tous n et k entiers, $\left|a_n^{(k)}\right| \leq b_n$. Montrer que, pour tout k entier, $\sum a_n^{(k)}$ converge et qu'on a $\lim_k \sum_{n=0}^{+\infty} a_n^{(k)} = \sum_{n=0}^{+\infty} a_n$.

14 - 40 Ⓢ ★★ Théorème de convergence monotone

Soit $(X_n)_{n\in\mathbf{N}^*}$ et X des variables aléatoires discrètes à valeurs réelles positives et telles que $\lim \uparrow X_n = X$ (au sens de la convergence simple).

a. Montrer que la suite $(\mathbf{E}(X_n))$ converge dans $\mathbf{R}_+ \cup \{+\infty\}$ et qu'on a $\lim \mathbf{E}(X_n) \leq \mathbf{E}(X)$.

b. Soit $(Y_n)_{n\in\mathbf{N}^*}$ et Z des variables aléatoires positives prenant un nombre fini de valeurs et telles que $\lim \uparrow Y_n = X$ et $Z \leq X$. Montrer $\lim \mathbf{E}(Y_n) \geq \mathbf{E}(Z)$. On pourra écrire $Y_n \geq (Z-\varepsilon)\mathbb{1}_{(Z-\varepsilon \leq Y_n)}$.

c. En déduire qu'il existe ℓ dans $\mathbf{R}_+ \cup \{+\infty\}$ tel que, pour toute suite $(Y_n)_{n\in\mathbf{N}^*}$ de variables aléatoires réelles prenant un nombre fini de valeurs et telles que $\lim \uparrow Y_n = X$, on ait $\lim \mathbf{E}(Y_n) = \ell$.

d. Montrer $\ell = \mathbf{E}(X)$.

e. Pour chaque entier n, on dispose d'une suite de variables aléatoires $(X_n^{(k)})_{k\in\mathbf{N}^*}$ prenant un nombre fini de valeurs, positives et telles que $\lim \uparrow X_n^{(k)} = X_n$. On pose alors $Y_k = \max(X_1^{(k)}, \ldots, X_k^{(k)})$.

i. Montrer que la suite (Y_k) est croissante et majorée par (X_k).

ii. On note Y la limite simple de (Y_k), à valeurs dans $\mathbf{R}_+ \cup \{+\infty\}$. Montrer que, pour tous entiers n et k avec $n \leq k$, $X_n^{(k)} \leq Y_k \leq X_k$ et en déduire $Y = X$.

iii. En déduire $\lim \mathbf{E}(Y_k) = \mathbf{E}(X)$ puis $\lim \uparrow \mathbf{E}(X_n) = \mathbf{E}(X)$.

14 - 41 Ⓢ ★★ Loi forte des grands nombres

Soit $(X_n)_{n\in\mathbf{N}^*}$ une suite de variables aléatoires discrètes réelles, indépendantes, centrées et admettant toutes un moment d'ordre 4 majoré par m. Montrer, pour n dans $\mathbf{N}^*$

$$\mathbf{E}\left((X_1 + \cdots + X_n)^4\right) \leq (3n^2 - 2n)m$$

et en déduire $\dfrac{X_1 + \cdots + X_n}{n} \xrightarrow{ps} 0$. On pourra s'inspirer de l'exercice 14 - 29.

14 - 42 Ⓢ ★★★ Convergence en loi

Soit $(X_n)_{n\in\mathbf{N}^*}$ une suite de variables aléatoires à valeurs dans $\mathbf{N}$, tout comme X. Montrer l'équivalence entre les deux assertions suivantes :

1. pour tout k dans $\mathbf{N}$, $\lim \mathbf{P}(X_n = k) = \mathbf{P}(X = k)$

2. $(G_{X_n})_{n\in\mathbf{N}^*}$ converge simplement vers G_X sur $[0;1[$.

Pour la réciproque, on utilisera un argument d'extraction diagonale pour démontrer que toutes les suites $(\mathbf{P}(X_n = k))_n$ ont une valeur d'adhérence et on montrera qu'en sus cette valeur d'adhérence est unique.

En déduire la convergence de la loi binomiale $\mathcal{B}(n,p)$, avec $np = \lambda$, vers la loi de Poisson $\mathcal{P}(\lambda)$.

14 - 43 Ⓢ ★★★★ Convergence dominée

Soit $(X_n)_{n\in\mathbf{N}^*}$ une suite de variables aléatoires discrètes à valeurs réelles et convergeant en probabilité vers une variable aléatoire discrète X. On suppose de plus qu'on dispose d'une variable aléatoire discrète réelle Z ayant un moment d'ordre 1 et telle que, pour tout entier n, $|X_n| \leq Z$.

a. On pose $Z_n = Z\mathbb{1}_{(Z\leq n)}$ et $Y_n = Z - Z_n$. Montrer $\lim \downarrow \mathbf{E}(Y_n) = 0$. On pourra utiliser le théorème de convergence monotone (exercice 14 - 40).

b. Soit $(B_k)_{k\in\mathbf{N}^*}$ une suite d'événements telle que $\lim \mathbf{P}(B_k) = 0$. Montrer $Z\mathbb{1}_{B_k} \leq n\mathbb{1}_{B_k\cap(Z\leq n)} + Z\mathbb{1}_{(Z>n)}$ et en déduire $\lim \mathbf{E}(Z\mathbb{1}_{B_k}) = 0$.

c. En déduire $\mathbf{P}(|X| \leq Z) = 1$. On pourra considérer les événements $(|X| > Z + 2^{-k})$.

d. En déduire $\lim \mathbf{E}(|X_n - X|) = 0$ puis $\lim \mathbf{E}(X_n) = \mathbf{E}(X)$. On pourra écrire $|X_n - X| = |X_n - X|\mathbb{1}_{(|X_n - X|\leq\varepsilon)} + |X_n - X|\mathbb{1}_{(|X_n-X|>\varepsilon)}$.