Continuité croissante et continuité décroissante

Théorème (continuité croissante). Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé. Soit $(\mathrm{A}_n)_{n\in\mathbb{N}}$ une suite d'événements de $\mathcal{A}$ croissante pour l'inclusion. On a alors l'égalité

$$\mathbb{P}\left(\bigcup_{n\in\mathbb{N}} \mathrm{A}_n\right) = \lim_{n\to+\infty} \mathbb{P}(\mathrm{A}_n).$$

Démonstration. Pour tout n dans $\mathbb{N}^*$, introduisons l'événement $\mathrm{B}_n = \mathrm{A}_n \setminus \mathrm{A}_{n-1}$. Posons également $\mathrm{B}_0 = \mathrm{A}_0$.

Fait 1. Les B_n sont deux à deux disjoints.

Démonstration du fait 1. Soit $n \in \mathbb{N}$. Soit un entier $j > n$. La définition de B_n et B_j donne

$$\mathrm{B}_n \subset \mathrm{A}_n \subset \mathrm{A}_{j-1} \qquad \text{et} \qquad \mathrm{B}_j = \mathrm{A}_j \setminus \mathrm{A}_{j-1} \qquad \text{donc} \qquad \mathrm{B}_j \cap \mathrm{B}_n = \varnothing.$$

Fait 2. La réunion des B_n est égale à la réunion des A_n.

Démonstration du fait 2. Soit $\omega \in \bigcup_{n\geqslant 0} \mathrm{A}_n$. L'ensemble des n tels que ω soit dans A_n est non vide donc il possède un plus petit élément, noté n_0.

Si n_0 vaut 0, alors ω appartient à A_0, c'est-à-dire à B_0, donc il appartient à la réunion de tous les B_n.

Si n_0 est strictement positif, alors ω appartient à A_{n_0} mais pas à A_{n_0-1} donc il appartient à B_{n_0}, donc il appartient à la réunion des B_n.

On a prouvé que la réunion des A_n est incluse dans la réunion des B_n.

Réciproquement, soit $\omega \in \bigcup_{n\geqslant 0} \mathrm{B}_n$. Il existe un entier n tel que ω appartienne à B_n. Que n soit nul ou non, la définition de B_n fait que B_n est inclus dans A_n donc ω est dans A_n.

Ainsi, la réunion des B_n est incluse dans la réunion des A_n.

Par double inclusion, ces deux réunions sont égales.

Ces deux faits prouvent que $\{\mathrm{B}_n\,;n\in\mathbb{N}\}$ est une partition (dénombrable) de la réunion des A_n. La σ-additivité donne donc

$$\mathbb{P}\left(\bigcup_{n\in\mathbb{N}} \mathrm{A}_n\right) = \sum_{n=0}^{+\infty} \mathbb{P}(\mathrm{B}_n).$$

Soit un entier N strictement positif. La somme partielle de rang $\mathrm{N}-1$ de la somme ci-dessus vaut

$$\sum_{n=0}^{\mathrm{N}-1} \mathbb{P}(\mathrm{B}_n) = \mathbb{P}(\mathrm{A}_0) + \sum_{n=1}^{\mathrm{N}-1} (\mathbb{P}(\mathrm{A}_n) - \mathbb{P}(\mathrm{A}_{n-1})) = \mathbb{P}(\mathrm{A}_0) + \mathbb{P}(\mathrm{A}_\mathrm{N}) - \mathbb{P}(\mathrm{A}_0) = \mathbb{P}(\mathrm{A}_\mathrm{N}).$$

La suite de terme général $\mathbb{P}(\mathrm{A}_\mathrm{N})$ est convergente car elle est croissante et majorée (par 1). On obtient donc

$$\lim_{\mathrm{N}\to+\infty} \mathbb{P}(\mathrm{A}_\mathrm{N}) = \sum_{n=0}^{+\infty} \mathbb{P}(\mathrm{B}_n) = \mathbb{P}\left(\bigcup_{n\in\mathbb{N}} \mathrm{A}_n\right).$$

La continuité croissante est prouvée. ♡

Théorème (continuité décroissante). Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé. Soit $(\mathrm{A}_n)_{n\in\mathbb{N}}$ une suite d'événements de $\mathcal{A}$ décroissante pour l'inclusion. On a alors l'égalité

$$\mathbb{P}\left(\bigcap \mathrm{A}_n\right) = \lim_{n\to+\infty} \mathbb{P}(\mathrm{A}_n).$$

Démonstration. La suite $(\overline{\mathrm{A}_n})_{n\geqslant 0}$ des événements contraires est croissante pour l'inclusion. La continuité croissante donne donc

$$\mathbb{P}\left(\bigcup \overline{\mathrm{A}_n}\right) = \lim_{n\to+\infty} \mathbb{P}(\overline{\mathrm{A}_n}).$$

L'événement contraire de la réunion des $\overline{\mathrm{A}_n}$ est l'intersection des A_n. On obtient donc

$$\mathbb{P}\left(\bigcap \mathrm{A}_n\right) = 1 - \mathbb{P}\left(\bigcup \overline{\mathrm{A}_n}\right) = 1 - \lim_{n\to+\infty} \mathbb{P}(\overline{\mathrm{A}_n}) = \lim_{n\to+\infty} (1 - \mathbb{P}(\overline{\mathrm{A}_n}) = \lim_{n\to+\infty} \mathbb{P}(\mathrm{A}_n).$$

La continuité décroissante est prouvée. ♡

Une propriété de l'espérance

Proposition. Soit X une variable aléatoire discrète à valeurs dans $\mathbb{N}$.

Alors l'espérance de X existe si, et seulement si, la série de terme général $\mathbb{P}(X \geqslant n)$ est convergente. De plus, si c'est le cas, alors l'espérance de X est donnée par

$$\mathbb{E}(X) = \sum_{n=1}^{+\infty} \mathbb{P}(X \geqslant n).$$

Démonstration. Soit N dans $\mathbb{N}^*$. On considère la somme partielle

$$\sum_{n=1}^{N} \mathbb{P}(X \geqslant n).$$

Pour chaque entier n strictement positif, on écrit $\mathbb{P}(X \geqslant n) = 1 - \mathbb{P}(X \leqslant n-1)$ car $[X \leqslant n-1]$ est l'événement contraire de $[X \geqslant n]$. Ensuite, on remarque que $[X \leqslant n-1]$ est la réunion disjointe des événements $[X = k]$ en faisant varier k de 0 à $n-1$. On obtient

$$\sum_{n=1}^{N} \mathbb{P}(X \geqslant n) = \sum_{n=1}^{N}(1 - \mathbb{P}(X \leqslant n-1)) = N - \sum_{n=1}^{N}\sum_{k=0}^{n-1} \mathbb{P}(X = k).$$

On permute les deux symboles de sommation

$$\sum_{n=1}^{N} \mathbb{P}(X \geqslant n) = N - \sum_{0 \leqslant k < n \leqslant N} \mathbb{P}(X = k) = N - \sum_{k=0}^{N-1}\sum_{n=k+1}^{N} \mathbb{P}(X = k) = N - \sum_{k=0}^{N-1}(N-k)\mathbb{P}(X = k).$$

On regroupe les termes qui ont N en facteur puis on reconnaît une somme partielle de la série en lien avec l'espérance de X.

$$\sum_{n=1}^{N} \mathbb{P}(X \geqslant n) = N(1 - \sum_{k=0}^{N-1} \mathbb{P}(X = k)) + \sum_{k=0}^{N-1} k\mathbb{P}(X = k),$$

donc

$$(*) \qquad \sum_{n=1}^{N} \mathbb{P}(X \geqslant n) = N\mathbb{P}(X \geqslant N) + \sum_{k=0}^{N-1} k\mathbb{P}(X = k).$$

À ce stade, on voit qu'il suffit de prouver que si l'une des deux séries converge, alors $N\mathbb{P}(X \geqslant N)$ tend vers 0 ; dans ce cas, on en déduit que l'autre série converge et que la somme est la même.

Première implication. On suppose que l'espérance de X existe. Prenons N dans $\mathbb{N}^*$ et remarquons la minoration

$$\sum_{k=N}^{+\infty} k\mathbb{P}(X = k) \geqslant \sum_{k=N}^{+\infty} N\mathbb{P}(X = k) = N\mathbb{P}(X \geqslant N).$$

Par ailleurs, le nombre $N\mathbb{P}(X \geqslant N)$ est bien sûr positif. On sait que la suite des restes d'une série convergente converge vers 0. Par le théorème des gendarmes, on en déduit que $N\mathbb{P}(X \geqslant N)$ tend vers 0 quand l'entier N tend vers $+\infty$.

D'après la relation $(*)$, la suite de terme général $\sum_{n=1}^{N} \mathbb{P}(X \geqslant n)$ converge vers $\mathbb{E}(X)$. Autrement dit, la série de terme général $\mathbb{P}(X \geqslant n)$ converge et sa somme vaut

$$\sum_{n=1}^{+\infty} \mathbb{P}(X \geqslant n) = \mathbb{E}(X).$$

Deuxième implication. On suppose maintenant que la série de terme général $\mathbb{P}(X \geqslant n)$ est convergente. Pour tout N dans $\mathbb{N}^*$, on trouve, en exploitant la relation $(*)$,

$$\sum_{k=0}^{N-1} k\mathbb{P}(X = k) = \sum_{n=1}^{N} \mathbb{P}(X \geqslant n) - N\mathbb{P}(X \geqslant N) \leqslant \sum_{n=1}^{N} \mathbb{P}(X \geqslant n) \leqslant \sum_{n=1}^{+\infty} \mathbb{P}(X \geqslant n).$$

Ainsi, la suite des sommes partielles de la série $\sum k\mathbb{P}(X = k)$ est majorée ; elle est également croissante (c'est une série à termes positifs) donc elle converge. On en déduit que l'espérance de X existe. On est donc ramené au cas de la première implication.

Lien entre la fonction génératrice et l'espérance

Proposition. Soit X une variable aléatoire discrète à valeurs dans $\mathbb{N}$.

Alors l'espérance de X existe si, et seulement si, la fonction génératrice G_X est dérivable en 0. Si c'est le cas, alors, l'espérance de X est donnée par

$$\mathbb{E}(X) = G'_X(1).$$

Démonstration. Dans un premier temps, on suppose que X est d'espérance finie. Autrement dit, la série $\sum n\mathbb{P}(X = n)$ est convergente.

Pour tout n dans $\mathbb{N}$, notons f_n la fonction $t \mapsto t^n \mathbb{P}(X = n)$.

$\boxed{1}$ On sait déjà que la série de fonctions $\sum\limits_{n\geqslant 0} f_n$ converge simplement sur $[0, 1]$. Sa somme est la fonction G_X.

$\boxed{2}$ Pour tout n dans $\mathbb{N}$, la fonction f_n est de classe $\mathcal{C}^1$ sur $[0, 1]$. Sa dérivée est la fonction nulle si $n = 0$ et c'est la fonction $t \mapsto nt^{n-1}\mathbb{P}(X = n)$ si $n \geqslant 1$.

$\boxed{3}$ Pour tout n dans $\mathbb{N}^*$, on trouve

$$||f'_n||_{\infty,\ [0,1]} = n\mathbb{P}(X = n).$$

Cette formule est encore valable pour $n = 0$. On en déduit que la série de terme général $||f'_n||_{\infty,\ [0,1]}$ est convergente. La série de fonctions $\sum f'_n$ converge normalement, donc uniformément, sur le segment $[0, 1]$.

Ces trois vérifications permettent d'affirmer que la fonction G_X est de classe $\mathcal{C}^1$ sur le segment $[0, 1]$ et que sa dérivée est donnée par

$$\forall t \in [0, 1], \qquad G'_X(t) = \sum_{n=1}^{+\infty} nt^{n-1}\mathbb{P}(X = n).$$

En particulier, on obtient

$$G'_X(1) = \sum_{n=1}^{+\infty} n\mathbb{P}(X = n) = \sum_{n=0}^{+\infty} n\mathbb{P}(X = n) = \mathbb{E}(X).$$

Passons à la réciproque. On suppose cette fois que la fonction G_X est dérivable en 1.

Rappelons que la fonction G_X est développable en série entière sur l'intervalle ouvert $]-1, 1[$, ce qui permet d'exprimer sa dérivée sur cet intervalle en dérivant terme à terme

$$\forall t \in]-1, 1[, \qquad G'_X(t) = \sum_{n=1}^{+\infty} \mathbb{P}(X = n)nt^{n-1}.$$

Cette formule permet de constater que la fonction G'_X est croissante sur l'intervalle $[0, 1[$. En particulier, cette fonction possède une limite en 1^-, qui est dans $\mathbb{R} \cup \{+\infty\}$.

Raisonnons par l'absurde et supposons que cette limite soit $+\infty$. On sait alors que la fonction G_X est continue sur le segment $[0, 1]$, de classe $\mathcal{C}^1$ sur $[0, 1[$ et que G'_X possède une limite infinie en 1^-. Le théorème de la limite de la dérivée permet alors d'en déduire que la fonction G_X n'est pas dérivable en 1.

Cette conclusion contredit l'hypothèse que G_X est dérivable en 1. La limite de G'_X en 1^- est donc une limite finie.

On sait donc que la fonction G_X est continue sur le segment $[0, 1]$, de classe $\mathcal{C}^1$ sur $[0, 1[$ que G'_X possède une limite finie en 1^-. On en déduit que la fonction G_X est de classe $\mathcal{C}^1$ sur $[0, 1]$.

Prenons un entier $N \geqslant 1$. Pour tout t dans $[0, 1[$, on connaît la majoration

$$\sum_{n=1}^{N} n\mathbb{P}(X = n)t^{n-1} \leqslant \sum_{n=1}^{+\infty} n\mathbb{P}(X = n)t^{n-1} = G'_X(t).$$

Faisons maintenant tendre t vers 1 par valeurs inférieures. La continuité de G'_X en 1 donne

$$\sum_{n=1}^{N} n\mathbb{P}(X = n) \leqslant G'_X(1).$$

Cette majoration est valable pour tout N dans $\mathbb{N}^*$. La suite des sommes partielles de la série $\sum n\mathbb{P}(\mathrm{X}=n)$ est donc majorée. Elle est par ailleurs croissante (les termes de cette série sont positifs) donc elle converge.

Ainsi, la série $\sum n\mathbb{P}(\mathrm{X}=n)$ est convergente. Autrement dit, la variable aléatoire X est d'espérance finie.

Le raisonnement de la première partie donne alors l'égalité $\mathrm{G}'_{\mathrm{X}}(1)=\mathbb{E}(\mathrm{X})$.

Fonction de répartition

Définition. Soit X une variable aléatoire réelle discrète définie sur un espace probabilisé $(\Omega,\mathcal{A},\mathbb{P})$.

La *fonction de répartition* de la variable aléatoire X est la fonction $\mathrm{F_X}:\mathbb{R}\to\mathbb{R}$ définie par

$$\mathrm{F_X}(x)=\mathbb{P}(\mathrm{X}\leqslant x).$$

Propriété. La fonction $\mathrm{F_X}$ admet pour limites 0 en $-\infty$ et 1 en $+\infty$.

Démonstration de la propriété. Pour tout n dans $\mathbb{N}$, notons $\mathrm{A}_n=[\mathrm{X}\leqslant n]$.

Pour tout n dans $\mathbb{N}$, l'inégalité $\mathrm{X}\leqslant n$ implique $\mathrm{X}\leqslant n+1$ donc A_n est inclus dans A_{n+1}.

La suite d'événements $(\mathrm{A}_n)_{n\in\mathbb{N}}$ est donc croissante pour l'inclusion, ce qui permettra d'appliquer la continuité croissante.

Fait. La réunion des A_n est égale à Ω.

Démonstration de ce fait. On sait déjà que la réunion des A_n est incluse dans Ω.
Réciproquement, soit $\omega\in\Omega$. Prenons $r=\lfloor\mathrm{X}(\omega)\rfloor+1$, de sorte que $r\geqslant\mathrm{X}(\omega)$.
Ainsi, l'élément ω réalise l'événement $[\mathrm{X}\leqslant r]$, c'est-à-dire A_r, si bien qu'il appartient à la réunion des A_n.
On a prouvé que Ω est inclus dans la réunion des A_n. Ces deux ensembles sont donc égaux.

La croissance de la suite d'événements $(\mathrm{A}_n)_{n\in\mathbb{N}}$ permet d'appliquer la continuité croissante, qui donne

$$\lim_{n\to+\infty}\mathbb{P}(\mathrm{A}_n)=\mathbb{P}\left(\bigcup_{n\in\mathbb{N}}\mathrm{A}_n\right),\qquad\text{c'est-à-dire}\qquad\lim_{n\to+\infty}\mathrm{F_X}(n)=\mathbb{P}(\Omega)=1.$$

Remarquons enfin que la fonction $\mathrm{F_X}$ est croissante, ce qui lui garantit une limite en $+\infty$. La suite $(\mathrm{F_X}(n))_{n\in\mathbb{N}}$ converge vers 1 donc la limite de la fonction $\mathrm{F_X}$ en $+\infty$ vaut 1.

Pour la limite en $-\infty$, on utilise la continuité décroissante avec les événements de la forme $\mathrm{B}_n=[\mathrm{X}\leqslant -n]$. ♡

Propriété (hors programme). Pour tout $x_0\in\mathbb{R}$, la fonction $\mathrm{F_X}$ est continue à droite en x_0.

Démonstration de la propriété. On suit le même raisonnement en appliquant la continuité décroissante aux événements de la forme $\mathrm{C}_n=[\mathrm{X}\leqslant x_0+2^{-n}]$. ♡

Inégalité de Cauchy-Schwarz pour la covariance

Théorème. Soient X et Y deux variables aléatoires admettant une variance. On a alors l'inégalité

$$\mathrm{Cov}(\mathrm{X},\mathrm{Y})^2\leqslant\mathbb{V}(\mathrm{X})\times\mathbb{V}(\mathrm{Y}).$$

De plus, l'égalité équivaut à ce que X suive une loi dégénérée ou à ce qu'il existe deux constantes a et b telles que $\mathrm{Y}-(a\mathrm{X}+b)$ soit presque sûrement nulle.

Démonstration du théorème. Si X suit une loi dégénérée, alors $\mathrm{Cov}(\mathrm{X},\mathrm{Y})$ et $\mathbb{V}(\mathrm{X})$ sont nuls donc l'égalité a lieu.

Dans la suite, on suppose que X ne suit pas une loi dégénérée, si bien que $\mathbb{V}(\mathrm{X})>0$.

On considère la fonction

$$f:t\mapsto t\mapsto\mathbb{V}(\mathrm{Y}-t\mathrm{X}),$$

définie de $\mathbb{R}$ dans $\mathbb{R}$. Le calcul donne

$$\forall t \in \mathbb{R}, \quad f(t) = t^2\mathbb{V}(\mathrm{X}) - 2t\,\mathrm{Cov}(\mathrm{X}, \mathrm{Y}) + \mathbb{V}(\mathrm{Y}).$$

La fonction f est polynomiale (à coefficients réels), de degré 2, à valeurs positives, donc son discriminant est négatif (au sens large), ce qui s'écrit

$$4\,\mathrm{Cov}(\mathrm{X}, \mathrm{Y})^2 - 4\mathbb{V}(\mathrm{X}) \times \mathbb{V}(\mathrm{Y}) \leqslant 0, \qquad \text{puis} \qquad \mathrm{Cov}(\mathrm{X}, \mathrm{Y})^2 \leqslant \mathbb{V}(\mathrm{X}) \times \mathbb{V}(\mathrm{Y}).$$

Passons au cas d'égalité.

Analyse. On suppose que l'égalité $\mathrm{Cov}(\mathrm{X}, \mathrm{Y})^2 = \mathbb{V}(\mathrm{X}) \times \mathbb{V}(\mathrm{Y})$ est réalisée. Le discriminant sus-mentionné est donc nul, si bien que la fonction f possède une racine réelle a — unique, mais ce n'est pas ce qui nous occupe. Cette racine est donnée par

$$a = \frac{\mathrm{Cov}(\mathrm{X}, \mathrm{Y})}{\mathbb{V}(\mathrm{X})}.$$

La variable aléatoire $\mathrm{Y} - a\mathrm{X}$ a alors une variance nulle, ce qui signifie qu'elle est presque sûrement égale à une constante b. Cette constante est égale à l'espérance de cette variable aléatoire, c'est-à-dire

$$b = \mathbb{E}(\mathrm{Y}) - a\mathbb{E}(\mathrm{X}).$$

La variable $\mathrm{Y} - (a\mathrm{X} + b)$ est donc presque sûrement nulle.

Synthèse. On suppose qu'il existe deux constantes a et b telles que $\mathrm{Y} - (a\mathrm{X} + b)$ soit presque sûrement nulle. On obtient alors que $\mathrm{Y} - a\mathrm{X}$ est de variance nulle, c'est-à-dire $f(a) = 0$.

Le discriminant de f est donc positif, donc nul, ce qui réalise le cas d'égalité dans l'inégalité de Cauchy-Schwarz. ♡