Probabilités discrètes

Tribu

Soit Ω un ensemble.

Définition

Une tribu sur Ω est une partie $\mathcal{F}$ de $\mathcal{P}(\Omega)$ vérifiant les trois axiomes suivants.

1. *Ω appartient à $\mathcal{F}$.*
2. *Si A appartient à $\mathcal{F}$, alors $\Omega \setminus A$ appartient à $\mathcal{F}$.*
3. *Si $(A_n)_{n\geqslant 0}$ est une suite d'éléments de $\mathcal{F}$, alors $\displaystyle\bigcup_{n=0}^{+\infty} A_n$ appartient à $\mathcal{F}$.*

Les éléments de $\mathcal{F}$ sont appelés **évènements**.

Probabilité

Définition

Une probabilité sur $(\Omega, \mathcal{F})$ *est une fonction* $\mathbb{P} : \mathcal{F} \to [0,1]$ *telle que :*

1. $\mathbb{P}(\Omega) = 1$
2. *Pour toute suite* $(A_n)_{n \geqslant 0}$ *d'éléments de* $\mathcal{F}$ *deux à deux disjoints et de réunion* A*,*
$$\mathbb{P}(A) = \sum_{n=0}^{+\infty} \mathbb{P}(A_n).$$

Le triplet $(\Omega, \mathcal{F}, \mathbb{P})$ *est appelé un* **espace probabilisé** *ou un* **espace de probabilité***.*

Conséquences faciles à vérifier en exercice

1) $\mathbb{P}(\emptyset) = 0$;
2) Pour tout événement A, $\mathbb{P}(\Omega \setminus A) = 1 - \mathbb{P}(A)$;
3) Pour tous événements A et B, si $A \subset B$, $\mathbb{P}(A) \leqslant \mathbb{P}(B)$;
4) Pour tous événements A et B, $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$;

Conséquences moins faciles très utiles

- Si $(A_n)_{n\geqslant 1}$ est une suite croissante d'événements de réunion A, $\mathbb{P}(A)$ est la limite de la suite croissante $(\mathbb{P}(A_n))_{n\geqslant 1}$;
- Si $(A_n)_{n\geqslant 1}$ est une suite décroissante d'événements d'intersection A, $\mathbb{P}(A)$ est la limite de la suite décroissante $(\mathbb{P}(A_n))_{n\geqslant 1}$.

Preuve : Soit $(A_n)_{n\geqslant 1}$ une suite croissante d'événements de réunion A. Soit $B_1 = A_1$ et pour $n \geq 2$, $B_n = A_n \setminus A_{n-1}$. La suite $(B_n)_{n\geqslant 1}$ est une suite d'éléments de $\mathcal{F}$ deux à deux disjoints et, pour tout $n \geqslant 1$, $A_n = \cup_{k=1}^n B_k$. Donc

$$\mathbb{P}(A_n) = \sum_{k=1}^{n} \mathbb{P}(B_k)$$

Cette série est donc convergente et de somme égale à

$$\sum_{k=1}^{+\infty} \mathbb{P}(B_k) = \mathbb{P}\left(\bigcup_{k=1}^{+\infty} B_k\right) = \mathbb{P}\left(\bigcup_{k=1}^{+\infty} A_k\right).$$

On obtient le résultat sur les suites décroissantes par passage au complémentaire.

Probabilités conditionnelles

Définition

Soit A un événement de probabilité $\mathbb{P}(A)$ non nulle. On définit une application $\mathbb{P}_A(\ ) : \mathcal{F} \to [0,1]$ par

$$\mathbb{P}_A(B) = \mathbb{P}(B \cap A)/\mathbb{P}(A).$$

La quantité $\mathbb{P}_A(B)$ est la probabilité conditionnelle de B sachant A.

Proposition

Soit A un événement de probabilité $\mathbb{P}(A)$ non nulle.
Alors $\mathbb{P}_A(\ ) : \mathcal{F} \to [0,1]$, $B \mapsto \mathbb{P}_A(B)$, est une probabilité sur $(\Omega, \mathcal{F})$.

Remarque

La relation précédente entraîne

$$\mathbb{P}(B \cap A) = \mathbb{P}(A)\mathbb{P}_A(B).$$

Probabilités conditionnelles, formules

Proposition (Probabilités conditionnelles en cascade)

$$\mathbb{P}(A_1 \cap \ldots \cap A_n) = \mathbb{P}(A_1)\mathbb{P}_{A_1}(A_2)\mathbb{P}_{A_1\cap A_2}(A_3)\cdots\mathbb{P}_{A_1\cap\cdots\cap A_{n-1}}(A_n),$$

à condition de donner au second membre la valeur zéro dès que l'un de ses termes est nul (même si certaines des probabilités conditionnelles y figurant ne sont alors pas définies).

Proposition (Formule des probabilités totales)

Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace de probabilité et $(A_i)_{i\in I}$ une partition dénombrable de Ω en événements de probabilités $\mathbb{P}(A_i)$ strictement positives pour tout $i \in I$. Alors, pour tout événement A,

$$\mathbb{P}(A) = \sum_{i\in I} \mathbb{P}_{A_i}(A)\mathbb{P}(A_i).$$

Probabilités conditionnelles, formule de Bayes

Proposition (Formule de Bayes)

Sous les hypothèses de la proposition précédente, pour tout événement A *de probabilité* $\mathbb{P}(A)$ *non nulle et pour tout* i,

$$\mathbb{P}_A(A_i) = \frac{\mathbb{P}_{A_i}(A)\mathbb{P}(A_i)}{\sum_{j \in I} \mathbb{P}_{A_j}(A)\mathbb{P}(A_j)}.$$

Indépendance d'évènements

Définition

Deux événements A et B sont indépendants si et seulement si

$$\mathbb{P}(B \cap A) = \mathbb{P}(A)\mathbb{P}(B).$$

Définition

Les événements $(A_i)_{i \in I}$ sont indépendants si et seulement si pour tout $J \subset I$, J fini non vide, on a

$$\mathbb{P}\left(\bigcap_{i \in J} A_i\right) = \prod_{i \in J} \mathbb{P}(A_i).$$

Exercice

Vérifier que si $|I| = 2$, les deux définitions coïncident.

Variables aléatoires discrètes

Définition

Soit $(\Omega, \mathcal{F}, \mathbb{P})$ *un espace de probabilité.*
Une variable aléatoire réelle est une application $X : \Omega \to \mathbb{R}$ *telle que, pour tout nombre réel* x*, l'ensemble* $\{\omega \in \Omega\,;\, X(\omega) \leqslant x\}$ *appartient à* $\mathcal{F}$*.*
Une variable aléatoire vectorielle est une application $X = (X_1, \dots, X_d) : \Omega \to \mathbb{R}^d$ *telle que, pour tout* i*,* X_i *est une variable aléatoire réelle.*
On dit que la variable aléatoire X *est discrète quand* $X(\Omega)$ *est dénombrable.*

Si $F \subset \mathbb{R}$, on notera $\{X \in F\}$ l'ensemble $\{\omega \in \Omega\,;\, X(\omega) \in F\}$.

Proposition

Si $f : \mathbb{R}^d \to \mathbb{R}^n$ *est continue et si* $X : \Omega \to \mathbb{R}^d$ *est une variable aléatoire, alors* $f \circ X$ *noté* $f(X)$ *est une variable aléatoire, à valeurs dans* $\mathbb{R}^n$*. Par exemple une somme et un produit de variables aléatoires sont des variables aléatoires.*

Loi d'une variable aléatoire discrète

Définition

Soit X une variable aléatoire discrète à valeurs dans $\mathbb{R}^d$ définie sur $(\Omega, \mathcal{F}, \mathbb{P})$. La loi de X (sous $\mathbb{P}$) est la probabilité sur $(\mathbb{R}^d, \mathcal{P}(\mathbb{R}^d))$, notée $\mathbb{P}_X$ et définie par la relation $\mathbb{P}_X(B) = \mathbb{P}(X \in B)$ pour tout $B \subset \mathbb{R}^d$.

On vérifie facilement que $\mathbb{P}_X$ est une probabilité :

- C'est une application à valeurs dans $[0, 1]$
- $\mathbb{P}_X(\mathbb{R}) = 1$
- Si $(B_n)_{n\in\mathbb{N}}$ est une suite de parties de $\mathbb{R}^d$ deux à deux disjointes, alors $\mathbb{P}_X(\cup_{n\in\mathbb{N}} B_n) = \sum_{n\in\mathbb{N}} \mathbb{P}_X(B_n)$.

Si $X(\Omega) = \{x_i\,;\, i \in I\}$ avec I dénombrable, la probabilité $\mathbb{P}_X$ est caractérisée par la donnée des nombres $\mathbb{P}_X(\{x_i\}) = \mathbb{P}(X = x_i)$.
En effet, pour $B \subset \mathbb{R}^d$, $\mathbb{P}_X(B) = \sum_{i\in I(B)} \mathbb{P}_X(\{x_i\})$ où $I(B)$ désigne l'ensemble des i dans I tels que x_i appartient à B, et cette somme est bien définie.

Loi d'un couple de variables aléatoires discrètes

Proposition (et définition)

Soient X et Y deux variables aléatoires réelles discrètes. On note $X(\Omega) = \{x_i\,;\, i \in I\}$ et $Y(\Omega) = \{y_j\,;\, j \in J\}$ avec I et J dénombrables. La loi de (X, Y) est donnée par $p_{i,j} = \mathbb{P}((X, Y) = (x_i, y_j))$. Alors les lois de X et Y sont appelées les lois marginales de (X, Y) et sont données par

$$\mathbb{P}_X(x_i) = \mathbb{P}(X = x_i) = \sum_{j \in J} p_{i,j}, \qquad \mathbb{P}_Y(y_j) = \mathbb{P}(Y = y_j) = \sum_{i \in I} p_{i,j}.$$

Fonction de répartition

Définition

Soit X une variable aléatoire réelle. On appelle fonction de répartition de X la fonction $F_X : \mathbb{R} \to [0,1]$ définie pour tout réel x par

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

Exercice

Dessiner la fonction de répartion d'une variable aléatoire constante, d'une variable aléatoire de loi de Bernoulli de paramètre p.

Propriétés des fonctions de répartition

Proposition

1) F_X est une fonction croissante et continue à droite.
2) La limite de F_X en $-\infty$ vaut 0.
3) La limite de F_X en $+\infty$ vaut 1.
4) Pour tout réel x, $F_X(x-) = \mathbb{P}(X < x)$.
5) Pour tout réel x, $\mathbb{P}(X = x) = F_X(x) - F_X(x-)$.

Le dernier point montre que la loi d'une variable aléatoire discrète est caractérisée par sa fonction de répartition.

Preuves

1) Si $x \leqslant y$, $\{X \leqslant x\} \subset \{X \leqslant y\}$. Par conséquent, $\mathbb{P}(X \leqslant x) \leqslant \mathbb{P}(X \leqslant y)$.

Pour le reste de la preuve, soit x un nombre réel, $(x_n)_{n \geqslant 0}$ une suite de nombres réels et $A_n := \{X \leqslant x_n\}$.

Si $(x_n)_{n \geqslant 0}$ décroît vers x, la suite $(A_n)_{n \geqslant 0}$ est décroissante et $\{X \leqslant x\}$ est l'intersection des A_n, donc $\mathbb{P}(X \leqslant x) = \lim_n \mathbb{P}(X \leqslant x_n)$.

2) Si $(x_n)_{n \geqslant 0}$ décroît vers $-\infty$, la suite $(A_n)_{n \geqslant 0}$ est décroissante et l'intersection des A_n est vide, donc $\lim_n \mathbb{P}(X \leqslant x_n) = 0$.

3) Si $(x_n)_{n \geqslant 0}$ croît vers $+\infty$, la suite $(A_n)_{n \geqslant 0}$ est croissante et la réunion des A_n est Ω tout entier, donc $\lim_n \mathbb{P}(X \leqslant x_n) = 1$.

4) Si $(x_n)_{n \geqslant 0}$ croît strictement vers x, la suite $(A_n)_{n \geqslant 0}$ est croissante et $\{X < x\}$ est la réunion des A_n, donc $\mathbb{P}(X < x) = \lim_n \mathbb{P}(X \leqslant x_n)$.

5) $\mathbb{P}(X = x) = \mathbb{P}(X \leqslant x) - \mathbb{P}(X < x)$.

Espérance d'une variable aléatoire réelle

Définition

Soit X *une variable aléatoire discrète réelle définie sur* $(\Omega, \mathcal{F}, \mathbb{P})$ *à valeurs dans* $X(\Omega) = \{x_i\,;\, i \in I\}$ *avec* $I \subset \mathbb{N}$ *. Pour tout* i *dans* I, *on note* $p_i = \mathbb{P}_X(\{x_i\}) = \mathbb{P}(X = x_i)$.
Si la famille $(x_i p_i)_{i \in I}$ *est sommable, on dit que* X *admet une espérance, notée* $E[X]$ *et définie par*

$$E[X] := \sum_{i \in I} x_i p_i.$$

Remarque

Si $X(\Omega)$ *est fini,* X *admet toujours une espérance.*
Si X *ne prend que des valeurs positives alors* $E[X] \geqslant 0$.
Si X *est une variable aléatoire constante égale à* x, *alors* $E[X] = x$.
Si a *et* b *sont des nombres réels et si* X *admet une espérance, alors* $aX + b$ *aussi et* $E[aX + b] = aE[X] + b$.

Espérance

- Espérance d'une fonction indicatrice :$E[\mathbf{1}_A] = \mathbb{P}(A)$.
- X admet une espérance si et seulement si $|X|$ admet une espérance et alors $|E[X]| \leq E[|X|]$.
- Si Y admet une espérance et si pour tout $\omega \in \Omega$, $|X(\omega)| \leq Y(\omega)$ alors X admet une espérance.

Proposition

Soit X et Y deux variables aléatoires réelles discrètes admettant des espérances. Alors $X + Y$ admet une espérance et

$$E[X + Y] = E[X] + E[Y].$$

Théorème de transfert

Proposition (Théorème de transfert)

Soit X une variable aléatoire discrète à valeurs dans $\mathbb{R}^d$ définie sur $(\Omega, \mathcal{F}, \mathbb{P})$. Notons $X(\Omega) = \{x_i\,;\, i \in I\}$ avec $I \subset \mathbb{N}$, l'ensemble des valeurs prises par X.
Soit $f : \mathbb{R}^d \to \mathbb{R}$ et notons $f(X(\Omega)) = \{y_j\,;\, j \in J\}$ avec $J \subset \mathbb{N}$.
Enfin, soit $Y = f(X)$. (i.e. $f \circ X$)
Si la famille $(f(x_i)\mathbb{P}(X = x_i))_{i \in I}$ est sommable, alors Y admet une espérance et

$$E[Y] = E[f(X)] = \sum_{i \in I} f(x_i)\mathbb{P}(X = x_i) = \sum_{j \in J} y_j\, \mathbb{P}_Y(\{y_j\}).$$

Définition

La variable aléatoire X est de carré intégrable si $E[X^2]$ est fini. On a alors par le théorème de transfert :

$$E[X^2] = \sum_{i \in I} x_i^2 \mathbb{P}(X = x_i).$$

Proposition (Inégalité de Cauchy-Schwarz)

Soit X et Y deux variables aléatoires réelles discrètes de carré intégrable. Alors la variable aléatoire XY admet une espérance et

$$|E[XY]| \leqslant E[X^2]^{1/2} E[Y^2]^{1/2}.$$

De plus, $X + Y$ est de carré intégrable.

Preuve de l'inégalité de Cauchy-Schwarz

Pour tout $\omega \in \Omega$, $2|X(\omega)Y(\omega)| \leqslant X^2(\omega) + Y^2(\omega)$. On en déduit par un résultat précédent que la variable aléatoire XY admet une espérance et que, pour tout réel t, la variable aléatoire $(X + tY)^2 = X^2 + 2tXY + t^2Y^2$ aussi. Comme $(X + tY)^2$ est positive partout, $E[(X + tY)^2] \geqslant 0$.
On développe et on traduit le fait que le polynome en t obtenu est toujours positif, ce qui nous donne le résultat.
Le discriminant du polynome est $\Delta = E[XY]^2 - E[X^2]E[Y^2]$.
On a donc égalité s'il existe t_0 tel que $E[(X + t_0Y)^2] = 0$, c'est à dire si X et Y sont proportionnelles.

Proposition (et définition)

Soit X une variable aléatoire discrète de carré intégrable. Alors X admet une espérance et

$$|E[X]| \leqslant E[|X|] \leqslant E[X^2]^{1/2}.$$

On définit alors la variance de X, notée $\mathrm{var}(X)$ ou $\sigma^2(X)$, et l'écart type de X, noté $\sigma(X)$, par

$$\mathrm{var}(X) = E[X^2] - E[X]^2 = E[(X - E[X])^2], \quad \sigma(X) = \sqrt{\mathrm{var}(X)}.$$

Pour tout réel x, $E[(X - x)^2] = \mathrm{var}(X) + (x - E[X])^2$ donc l'espérance est aussi la valeur de x qui minimise $E[(X - x)^2]$, et la valeur minimale obtenue est la variance.

Preuve : La première inégalité est une conséquence de la proposition précédente appliquée aux variables aléatoires X et $Y = 1$. Par ailleurs,

$$E[(X - x)^2] = E[X^2] - 2xE[X] + x^2 = E[X^2] - E[X]^2 + (E[X] - x)^2,$$

ce qui termine la preuve

Ecart type

L'écart type mesure la déviation autour de l'espérance.
Si la variable aléatoire représente une quantité mesurable (par exemple une longueur, une température, etc ...), l'écart type a la même dimension et non pas la variance.
Soit X une variable aléatoire ayant une espérance m et un écart type σ.
Alors la variable aléatoire X^* définie par $X^* = \frac{X-m}{\sigma}$ est **centrée réduite**, c'est à dire que son espérance est nulle et sa variance est 1. De plus X^* est sans dimension.

Covariance de deux variables aléatoires

Proposition (et définition)

Soient X et Y deux variables aléatoires réelles discrètes de carré intégrable. On définit la covariance de X et Y, notée $cov(X, Y)$, par

$$cov(X, Y) = E[(X - E[X])(Y - E[Y])] = E[XY] - E[X]E[Y].$$

On a alors

$$\mathrm{var}(X + Y) = \mathrm{var}(X) + \mathrm{var}(Y) + 2cov(X, Y).$$

Exercice

Faire la preuve. Quelle est la formule donnant la variance d'une somme de n variables aléatoires en fonction de leurs variances et covariances ?

Remarque

Une variance est toujours positive ou nulle mais une covariance peut être positive ou nulle ou négative.

Variables aléatoires indépendantes

Définition

Les variables aléatoires discrètes X_k, $1 \leqslant k \leqslant n$, à valeurs dans $E_1,\dots,E_n$ respectivement, sont indépendantes si et seulement si, pour tout $(x_1,\dots,x_n)$ dans $E_1 \times \dots \times E_n$,

$$\mathbb{P}(\forall 1 \leqslant k \leqslant n,\ X_k = x_k) = \prod_{k=1}^{n} \mathbb{P}(X_k = x_k).$$

Définition

Une suite de variables aléatoires discrètes $(X_n)_{n\geq 1}$ est dite indépendante si pour tout $n \geq 1$, les variables aléatoires X_k, $1 \leq k \leq n$ sont indépendantes.

Indépendance et calcul d'espérance

Proposition

Les variables aléatoires discrètes $(X_1, \dots, X_n)$ *à valeurs respectivement dans* $E_1, \dots, E_n$ *sont indépendantes si et seulement si pour toutes fonctions* $f_k : E_k \to \mathbb{R}$ *positives (resp. bornées),*

$$E\left[\prod_{k=1}^{n} f_k(X_k)\right] = \prod_{k=1}^{n} E[f_k(X_k)].$$

Dans ce cas, si on se donne, pour tout k*, une fonction* $g_k : E_k \to \mathbb{R}$ *telle que* $g_k(X_k)$ *est intégrable et si on pose* $Z := g_1(X_1) \cdots g_n(X_n)$*, alors* Z *est intégrable et*

$$E[Z] = \prod_{k=1}^{n} E[g_k(X_k)].$$

Indépendance et covariance

Proposition

Soit X et Y deux variables aléatoires réelles discrètes de carré intégrable. On suppose que X et Y sont indépendantes. Alors,

$$E[XY] = E[X]E[Y], \quad cov(X, Y) = 0, \quad \mathrm{var}(X + Y) = \mathrm{var}(X) + \mathrm{var}(Y).$$

Contre exemple

Attention : la réciproque au résultat précédent est fausse, on peut trouver des variables aléatoires X et Y telles que $\text{cov}(X, Y) = 0$ sans que X et Y soient indépendantes.
Par exemple, soit X une variable aléatoire de loi $\mathbb{P}(X = 1) = \mathbb{P}(X = -1) = \frac{1}{2}$, Z une variable aléatoire indépendante de X et de loi $\mathbb{P}(Z = -1) = \frac{2}{3}$ et $\mathbb{P}(Z = 2) = \frac{1}{3}$, $Y = XZ$.
Montrer que $E[X] = E[Y] = E[XY] = 0$ donc X et Y ne sont pas corrélées, mais que $\mathbb{P}(X = 1,\ Y = -2) = 0$, et en conclure que X et Y ne sont pas indépendantes.

Fonctions génératrices

Définition

Soit X une variable aléatoire à valeurs dans $\mathbb{N} \cup \{+\infty\}$. La fonction génératrice de X est la fonction $g_X : [0,1] \to [0,1]$ définie par

$$g_X(s) := \sum_{n=0}^{+\infty} s^n \mathbb{P}(X = n) = E[s^X].$$

Propriétés des fonctions génératrices

Proposition

Toute fonction génératrice vérifie les propriétés suivantes.

1. *La fonction g_X est convexe, croissante sur $[0,1]$ et indéfiniment dérivable sur $[0,1[$.*
2. *La fonction g_X détermine la loi de X.*
3. *Si on suppose que X est à valeurs dans $\mathbb{N}$, X admet un moment d'ordre p si et seulement si g_X admet une dérivée à gauche d'ordre p en 1. Dans ce cas, pour $p = 1$, $E[X] = g_X'(1-)$ et pour $p \geqslant 1$,*

$$E[X(X-1)\cdots(X-p+1)] = g_X^{(p)}(1-).$$

4. *Si X et Y sont indépendantes, alors $g_{X+Y} = g_X g_Y$.*

Preuve

La fonction g_X est donnée par une série entière de rayon de convergence au moins égal à 1, car $\sum_n \mathbb{P}(X = n)$ converge. C'est donc une fonction indéfiniment dérivable sur $[0, 1[$, et

$$g_X^{(p)}(s) = \sum_{n=0}^{+\infty} n(n-1)\cdots(n-p+1)s^{n-p}\mathbb{P}(X = n),$$

donc

$$g_X^{(p)}(s) = E[X(X-1)\cdots(X-p+1)s^{X-p}].$$

En particulier, $g_X^{(n)}(0) = n!\,\mathbb{P}(X = n)$. On en déduit facilement les points 1. et 2.

Preuve, suite

Démontrons le point 3. pour $p = 1$. On suppose donc que X est à valeurs dans $\mathbb{N}$ et on pose $g = g_X$. Alors, $h(s) := \frac{g(1)-g(s)}{1-s} = \sum_{n=0}^{+\infty} \frac{1-s^n}{1-s} \mathbb{P}(X = n)$.

Supposons que X admet une espérance. On va utiliser le fait que, pour tout entier $n \geqslant 0$ et pour tout s dans $[0, 1[$, $0 \leqslant \frac{1-s^n}{1-s} \leqslant n$.

Par conséquent, $h(s) \leqslant \sum_{n=0}^{+\infty} n\mathbb{P}(X = n) = E[X]$.

Dans l'autre sens, pour tout entier $N \geqslant 0$,

$$h(s) \geqslant \sum_{n=0}^{N} \frac{1-s^n}{1-s} \mathbb{P}(X = n),$$

et la fraction $(1 - s^n)/(1 - s)$ tend vers n quand s tend vers 1. Pour N fixé, on ne manipule que des sommes finies, donc

$$\lim_{s \to 1} h(s) \geqslant \sum_{n=0}^{N} n\mathbb{P}(X = n).$$

Preuve, suite

On a le droit d'écrire ceci car h étant croissante sur $[0, 1[$ admet une limite quand s tend vers 1 par valeurs inférieures. Cette minoration est vraie pour tout $N \geqslant 0$, donc

$$\lim_{s \to 1} h(s) \geqslant \sup_N \sum_{n=0}^{N} n\mathbb{P}(X = n) = \sum_{n=0}^{+\infty} n\mathbb{P}(X = n) = E[X],$$

donc $h(s)$ tend vers $E[X]$ quand s tend vers 1, c'est-à-dire que la dérivée à gauche en 1 de g existe et vaut $E[X]$,

Preuve, suite

Réciproquement, supposons que g admet une dérivée à gauche en 1. Alors, pour tout $N \geqslant 0$,

$$\sum_{n=0}^{N} n\mathbb{P}(X = n) = \lim_{s \to 1} \sum_{n=0}^{N} \frac{1 - s^n}{1 - s} \mathbb{P}(X = n) \leqslant \lim_{s \to 1} \frac{1 - g(s)}{1 - s} = g'(1-),$$

donc X admet une espérance et peut on utiliser la partie précédente de la preuve pour conclure.

Finalement, g est dérivable à gauche en 1 si et seulement si $\sum_n n\mathbb{P}(X = n)$ converge si et seulement si X admet une espérance.

Le résultat général se prouve de la même manière.

Preuve, suite

Pour le point 4., pour tout s dans $[0,1]$,

$$E[s^{X+Y}] = E[s^X s^Y] = E[s^X]E[s^Y],$$

grâce à une conséquence de l'indépendance et au fait que la fonction de $\mathbb{N}$ dans $\mathbb{R}$ définie par $x \mapsto s^x$ est bornée pour tout s dans $[0,1]$.

Remarque

Attention, la réciproque du point 4 est fausse.

Soit (X, Y) dont la loi est donnée dans le tableau suivant :

$x \backslash y$	0	1	2	$\mathbb{P}(X = x)$
0	$1/9$	0	$2/9$	$1/3$
1	$2/9$	$1/9$	0	$1/3$
2	0	$2/9$	$1/9$	$1/3$
$\mathbb{P}(Y = y)$	$1/3$	$1/3$	$1/3$	

Montrer que X et Y ne sont pas indépendantes mais que $g_{X+Y} = g_X g_Y$.

Lois à connaître, loi de Bernoulli $b(p)$

Paramètre p avec $0 \leqslant p \leqslant 1$: X à valeurs dans $\{0,1\}$ et $\mathbb{P}(X=1) = p = 1 - \mathbb{P}(X=0)$.

Donc $E[X] = p$, $\mathrm{var}(X) = p(1-p)$, et $g_X(s) = 1 - p + ps$.

Modèle : jeu de pile ou face.

Lois à connaître, loi uniforme $U(A)$

Paramètre A un ensemble fini de cardinal $n \geqslant 1$: X est à valeurs dans A et pour tout x dans A, $\mathbb{P}(X = x) = 1/n$.

Modèle : expérience avec n résultats possibles, si on ne dispose d'aucune information supplémentaire, ou bien si la situation est entièrement symétrique.

Lois à connaître, loi binomiale $B(n, p)$

Paramètres (n, p) avec $0 \leqslant p \leqslant 1$ réel et $n \geqslant 0$ entier : X est à valeurs dans $\{0, 1, \dots, n\}$ et, pour tout $0 \leqslant k \leqslant n$, $\mathbb{P}(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$.
Donc $g_X(s) = (1 - p + ps)^n$.
Modèle : nombre de piles parmi n résultats d'un jeu de pile ou face.

Proposition

Soit $(X_1, \dots, X_n)$ n variables aléatoires de Bernoulli indépendantes de paramètre p. Alors $X = X_1 + \cdots + X_n$ suit une loi binomiale de paramètre (n, p).

Lois à connaître, loi binomiale $B(n, p)$ suite

Preuve : On peut utiliser les fonctions génératrices. Grâce à l'indépendance,

$$g_X(s) = \prod_k g_{X_k}(s) = (1 - p + ps)^n.$$

Comme la fonction génératrice caractérise la loi, le résultat suit.
Donc $E[X] = np$ et $\mathrm{var}(X) = np(1-p)$.

Proposition

Si n tend vers l'infini et p tend vers 0 de sorte que np tende vers a avec a strictement positif et fini, alors, pour tout $k \geqslant 0$,

$$\mathbb{P}(X_n = k) \to \frac{a^k}{k!}\mathrm{e}^{-a}.$$

Démontrer cette proposition.

Lois à connaître, loi de Poisson $P(a)$

Paramètre a strictement positif : X à valeurs dans $\mathbb{N}$ et $\mathbb{P}(X = k) = \dfrac{a^k}{k!}\mathrm{e}^{-a}$ pour tout $k \geqslant 0$.
Donc $g_X(s) = \mathrm{e}^{-a(1-s)}$, $E[X] = a$ et $\mathrm{var}(X) = a$.
Modèle : d'après la proposition précédente, loi des événements rares.

Lois à connaître, loi géométrique $G(p)$

Paramètre p dans $[0,1]$: X est à valeurs dans $\mathbb{N}$ et $\mathbb{P}(X=n)=p(1-p)^n$ pour tout $n\geqslant 0$.
Donc $g_X(s)=p/(1-(1-p)s)$, $E[X]=(1-p)/p$ et $\mathrm{var}(X)=(1-p)/p^2$.
Modèle : instant du premier succès dans une suite d'expériences indépendantes chacune avec probabilité p de succès.

Proposition

Soit $(X_n)_{n\geqslant 0}$ une suite de variables aléatoires indépendantes de loi de Bernoulli de paramètre p (c'est-à-dire que pour tout $n\geqslant 0$, les variables aléatoires $(X_0,\ldots,X_n)$ sont indépendantes). Soit X la variable aléatoire définie par $X=\min\{n\geqslant 0\,|\,X_n=1\}$.
Alors X est finie avec probabilité 1 et de loi géométrique de paramètre p.

Preuve : L'événement $\{X=n\}$ vaut l'intersection de $\{X_n=1\}$ et des n événements $\{X_k=0\}$ pour $0\leqslant k\leqslant n-1$. Par indépendance, $\mathbb{P}(X=n)=p(1-p)^n$.

Lois à connaître, loi hypergéométrique $H(n, N, N_1)$

Paramètres (n, N, N_1) avec $0 \leqslant n \leqslant N$ et $0 \leqslant N_1 \leqslant N$: X est à valeurs dans $\{0, 1, \dots, n\}$ et pour tout $0 \leqslant k \leqslant n$,

$$\mathbb{P}(X = k) = \frac{\binom{N_1}{k}\binom{N-N_1}{n-k}}{\binom{N}{n}},$$

avec la convention que $\binom{i}{j} = 0$ si $j < 0$ ou si $j > i$.

Modèle : on effectue n tirages sans remise dans un ensemble A de cardinal N ; le nombre d'objets de $A_1 \subset A$ tirés suit cette loi si le cardinal de A est N et celui de A_1 est N_1.

Proposition

Soit X une variable aléatoire de loi hypergéométrique $H(n, N, N_1)$. Alors,

$$E[X] = nN_1/N, \qquad \mathrm{var}(X) = \frac{nN_1(N - N_1)(N - n)}{N^2(N - 1)}.$$

Lois à connaître, loi hypergéométrique $H(n, N, N_1)$, suite

Preuve : On utilise le fait que $k\binom{N_1}{k} = N_1\binom{N_1-1}{k-1}$. Ainsi,

$$k\,\mathbb{P}(X=k) = \frac{k\binom{N_1}{k}\binom{N-N_1}{n-k}}{\binom{N}{n}} = \frac{N_1\binom{N_1-1}{k-1}\binom{N-N_1}{(n-1)-(k-1)}}{N\binom{N-1}{n-1}/n},$$

donc

$$k\,\mathbb{P}(X=k) = (nN_1/N)\mathbb{P}(Y=k-1),$$

si Y suit une loi hypergéométrique de paramètres $(n-1, N-1, N_1-1)$.
En sommant sur k et en utilisant le fait que la somme des $\mathbb{P}(Y=k-1)$ vaut 1, on trouve la valeur de $E[X]$.
En appliquant le même principe à $\mathbb{P}(Y=k-1)$, on obtient

$$k(k-1)\,\mathbb{P}(X=k) = (nN_1/N)(k-1)\mathbb{P}(Y=k-1),$$

D'où $E[X(X-1)] = \frac{n(n-1)N_1(N_1-1)}{N(N-1)}$
ce qui permet de trouver la valeur de
$\mathrm{var}(X) = E[X(X-1)] + E[X] - E[X]^2$.
Fin de la preuve.

Lois à connaître, loi hypergéométrique $H(n, N, N_1)$, suite

Proposition

Si N et N_1 tendent vers l'infini de sorte que N_1/N converge vers p, alors $\mathbb{P}(X = k) \to \binom{n}{k} p^k (1-p)^{n-k}$.

Preuve laissée en exercice.
Ceci montre qu'un tirage sans remise équivaut à un tirage avec remise si le cardinal de l'ensemble est grand et le rapport N_1/N à peu près constant.

Remarque

Soit X_i la variable aléatoire valant 1 si au tirage numéro i, on obtient un objet de A_1 et 0 sinon. Alors $X = X_1 + \cdots + X_n$ et les variables aléatoires X_i suivent des lois de Bernoulli. Quels sont leurs paramètres ? Puisque $E[X] = nN_1/N$ et $E[X] = E[X_1] + \cdots + E[X_n]$, on voit que $\mathbb{P}(X_i = 1) = N_1/N$ pour tout i.
Mais attention, les variables aléatoires X_i ne sont pas indépendantes ; sinon la loi de leur somme X serait binomiale $(n, N_1/N)$.