Leçon 264 : Variables aléatoires discrètes. Exemples et applications.

RM
2022-2023

On se place dans un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ et $(E, \mathcal{E})$ un espace probabilisable.

1 Variables aléatoires discrètes

1.1 Définitions

Définition 1 : On appelle variable aléatoire de Ω dans E une application $X : \Omega \to E$ mesurable, c'est-à-dire que $\forall F \in \mathcal{E}, X^{-1}(F) \in \mathcal{A}$. La fonction $\mathbb{P}_X : \mathcal{E} \to [0,1], F \mapsto \mathbb{P}(X^{-1}(F))$ est alors une probabilité sur $(E, \mathcal{E})$, appelée loi de probabilité de X.

Remarque 2 : On choisit très souvent $E = \mathbb{N}$ ou $E = \mathbb{R}$ muni respectivement de la tribu discrète ou de la tribu borélienne.

Définition 3 : On dit qu'une variable aléatoire $X : \Omega \to E$ est discrète si $X(\Omega)$ est fini ou dénombrable.

Remarque/Exemple 4 : L'espace Ω n'est pas forcément fini ou dénombrable. Par exemple l'espace Ω des suites $\omega = (\omega_n)_{n\in\mathbb{N}}$ à valeurs dans $\{0,1\}$ est infini non dénombrable, et la variable aléatoire X tel que $X(\omega) = \inf\{n \in \mathbb{N} | \omega_n = 0\}$ est discrète car à valeurs dans $\mathbb{N} \cup \{+\infty\}$.

Exemple 5 : Soit un jeu de 52 cartes et X la variable aléatoire qui associe 1 si la carte tiré est noir et 0 si elle rouge. Alors X est discrète.

Proposition 6 : Une variable aléatoire discrète X est caractérisée par les valeurs $\mathbb{P}(X = x)$ avec $x \in X(\Omega)$. Ainsi pour tout $A \subset X(\Omega)$, on a $\mathbb{P}(X \in A) = \sum_{x\in A} \mathbb{P}(X = x)$.

Remarque 7 : Plus généralement, si on note $(x_i)_{i\in I}$ ou I est donc au plus dénombrable tel que $\mathbb{P}(X = x_i) \neq 0$, alors on a que la loi de X est

$$\mathbb{P}_X = \sum_{i\in I} \mathbb{P}(X = x_i)\delta_{x_i}$$

avec évidemment $\sum_{i\in I} \mathbb{P}(X = x_i) = 1$.

1.2 Lois discrètes usuelles

Définition 8 : On définit en annexe les variables aléatoires discrètes les plus classiques, qui sont à valeurs dans $\mathbb{N}$, avec leur loi, espérance, variance... .

Application (modélisation par des variables aléatoires discrètes) 9 : Pour celles usuelles :

Uniforme : Le résultat d'un dé à 6 faces jeté suit la loi uniforme $\mathcal{U}([\![1\,;6]\!])$.

Bernoulli : Le résultat d'un pile ou face en lançant une pièce suit une loi de Bernoulli de paramètre $\frac{1}{2}$.

Binomial : Si on lance n fois la pièce de manière indépendante, alors le nombre de succès (pile ou face) suit une loi binomial de paramètres n et $\frac{1}{2}$.

Poisson : Si il y a une probabilité p qu'un client rentre chaque seconde dans un magasin, alors le nombre de client qui entrent sur un intervalle d'une heure suit approximativement une loi de poisson de paramètre $\lambda = 3600p$.

Géométrique : Le nombre de lancer à effectuer d'une pièce équilibré avant d'obtenir un face suit une loi géométrique de paramètre $\frac{1}{2}$.

Proposition 10 : Si $X \sim \mathcal{B}(n,p)$ et $Y \sim \mathcal{B}(m,p)$ indépendante, alors $X + Y \sim \mathcal{B}(n + m, p)$.

Proposition 11 : Une loi de poisson est la limite d'une loi binomial de paramètre $(n, \lambda/n)$ lorsque $n \to +\infty$.

Proposition 12 : Soit $X \sim \mathcal{P}(\lambda)$ et $Y \sim \mathcal{P}(\mu)$ indépendantes, alors $X+Y \sim \mathcal{P}(\lambda+\mu)$.

Proposition 13 : Soit X une variable aléatoire à valeur dans $\mathbb{N}^*$. Alors X suit une loi géométrique si et seulement si $\mathbb{P}(X = 1) \in]0,1[$ et X est sans mémoire, c'est-à-dire que pour tout $(n,k) \in (\mathbb{N}^*)^2$, $\mathbb{P}(X > n + k | X > n) = \mathbb{P}(X > n)$.

1.3 Indépendance et conditionnement

Définition (Probabilité conditionnelle) 14 : Soit B un événement, avec $\mathbb{P}(B) > 0$. Soit $A \in \mathcal{A}$ un événement, alors on définit la probabilité conditionnelle de A sachant B par $\mathbb{P}(A|B) = \frac{\mathbb{P}(A\cap B}{\mathbb{P}(B)}$. L'application $\mathbb{P}_B : \mathcal{A} \to [0,1], A \mapsto \mathbb{P}(A \cap B)/\mathbb{P}(B)$ définit une nouvelle probabilité sur $(\Omega, \mathcal{A})$, appelé probabilité conditionnelle sachant B.

Proposition 15 : Soient $A_1, ..., A_n$ un nombre fini d'événements tel que $\mathbb{P}(\cap_{i=1}^{n-1} A_i) > 0$. Alors $\mathbb{P}(\cap_{i=1}^{n} A_i) = \mathbb{P}(A_1)\mathbb{P}(A_2|A_1)\mathbb{P}(A_3|A_1 \cap A_2)...\mathbb{P}(A_n|A_1 \cap ... \cap A_{n-1})$.

Exemple 16 : La probabilité de tirée 4 as sans remise dans un jeu de 52 cartes est de $1/52 * 51 * 50 * 49 = 1/6497000$ (ce qui est donc très peu probable).

Proposition (Formule des probabilités totales) 17 : Soit $(A_i)_{i\in I}$ une famille finie ou dénombrable d'événements formant une partition de Ω, avec $\mathbb{P}(A_i) > 0$ pour tout $i \in I$. Alors pour tout $A \in \mathcal{A}$, on a $\mathbb{P}(A) = \sum_{i\in I} \mathbb{P}(A|A_i)\mathbb{P}(A_i)$.

Corollaire (Formule de Bayes) 18 : Sous les même hypothèses avec $\mathbb{P}(A) > 0$, on a pour tout $i \in I$, $\mathbb{P}(A_i|A) = \dfrac{\mathbb{P}(A|A_i)\mathbb{P}(A_i)}{\sum_{i\in I}\mathbb{P}(A|A_i)\mathbb{P}(A_i)}$.

Exemple (Paradoxe du faux positif) 19 : Soit un test de détection d'une maladie rare positif à 99% lorsqu'un individu est malade, et positif à $0,1\%$ quand il n'est pas malade. Si $0,01\%$ est malade, alors la probabilité d'être malade si le test est positif est de $0,09$, soit 9%, ce qui est semble assez paradoxal.

Définition 20 : Deux événements A et B sont dit indépendant si $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$.

Définition 21 : Soit X et Y deux variables aléatoires discrètes de Ω vers E. Soit $y \in E$ tel que $\mathbb{P}(Y = y) \neq 0$. On appelle loi conditionnelle de X sachant $(Y = y)$ la probabilité $\mathbb{P}_y$ définie sur $X(\Omega)$ par $\mathbb{P}_y : x \in X(\Omega) \mapsto \mathbb{P}(X = x|Y = y) = \dfrac{\mathbb{P}(X = x \cap Y = y)}{\mathbb{P}(Y = y)}$.

Définition 22 : Deux variables aléatoires discrètes X et Y de Ω vers E sont indépendantes si et seulement si $\forall(x, y) \in X(\Omega) \times Y(\Omega)$, $\mathbb{P}(X = x \cap Y = y) = \mathbb{P}(X = x)\mathbb{P}(Y = y)$.
Les variables aléatoires discrètes $(X_i)_{i\in I}$ sont mutuellement indépendantes si et seulement si pour toute partie finie $J \subset I$, on a $\forall(x_j)_{j\in J} \in \prod_{j\in J} X_j(\Omega), \mathbb{P}(\cap_{j\in J}X_j = x_j) = \prod_{j\in J}\mathbb{P}(X_j = x_j)$.

2 De l'espérance aux fonctions génératrices

2.1 Espérance, variance

Définition (Espérance) 23 : Soit X une variable aléatoire discrète réelle sur Ω. On dit que X admet une espérance si la famille $(x\mathbb{P}(X = x))_{x\in X(\Omega)}$ est sommable, ou de manière équivalente, si $\sum_{x\in X(\Omega)} |x|\mathbb{P}(X = x) < +\infty$. On appelle alors espérance de X la valeur notée $E(X)$ définie par $E(X) = \sum_{x\in X(\Omega)} x\mathbb{P}(X = x)$.

Remarque 24 : Une variable aléatoire finie admet toujours une espérance. On dit que X est centré si $E(X) = 0$.

Proposition 25 : Soient X et Y deux variables aléatoires discrètes réelles sur Ω admettant une espérance et $\lambda \in \mathbb{R}$. Alors
$i)$ $X + Y$ et λX admettent une espérance avec $E(X + Y) = E(X) + E(Y)$ et $E(\lambda X) = \lambda X$.
$ii)$ si $X \geq 0$, alors $E(X) \geq 0$.
$iii)$ on a $|E(X)| \leq E(|X|)$.
$iv)$ Si X est Y sont indépendantes, alors XY admet une espérance et $E(XY) = E(X)E(Y)$.

Théorème (de transfert) 26 : Soit X une variable aléatoire discrète de Ω sur E et $f : E \to \mathbb{R}$. La variable aléatoire discrète réelle $f(X)$ admet une espérance si et seulement si la famille $(f(x)\mathbb{P}(X = x))_{x\in X(\Omega)}$ est sommable et alors $E(f(X)) = \sum_{x\in X(\Omega)} f(x)\mathbb{P}(X = x)$.

Proposition 27 : Soit $X : \Omega \to \mathbb{N}$ une variable aléatoire discrète. Alors X admet une espérance si et seulement si la famille $(\mathbb{P}(X > n))_{n\in\mathbb{N}}$ est sommable, et on a $E(X) = \sum_{n=0}^{+\infty} \mathbb{P}(X > n)$.

Définition 28 : Soit $p \in \mathbb{N}^*$ et X une variable aléatoire discrète réelle sur Ω. On dit que X admet un moment d'ordre p lorsque X^p admet une espérance. On appelle alors moment d'ordre p de X la valeur $E(X^p) = \sum_{x\in X(\Omega)} x^p\mathbb{P}(X = x)$.

Définition (Variance) 29 : Soit X une variable aléatoire discrète réelle su Ω, admettant un moment d'ordre 2. On appelle variance de X la valeur $V(X)$ définie par $V(X) = E((X - E(X))^2)$. On appelle écart type la valeur $\sigma(X) = \sqrt{V(X)}$.

Remarque 30 : On dit que X est réduite si $V(X) = 1$. Pour toute variable aléatoire réelle discrète X de variance non nulle, la variable aléatoire $X^* = (X - E(X))/\sigma(X)$ est centrée réduite.

Proposition 31 : Soit X une variable aléatoire discrète réelle, admettant un moment d'ordre 2. Alors
$i)$ On a $V(X) = E(X^2) - E(X)^2$ (Köenig-Huygens).
$ii)$ Pour tout $a, b \in \mathbb{R}$, $V(aX + b) = a^2V(X)$.

Définition 32 : Soient X et Y deux variables aléatoires discrètes réelles, admettant un moment d'ordre 2. On apelle covariance de X et Y la valeur notée $Cov(X, Y)$ définie par $Cov(X, Y) = E((X - E(X))(Y - E(Y)))$.

Proposition 33 : Soient X et Y deux variables aléatoires discrètes réelles, admettant un moment d'ordre 2. Alors
$i)$ On a $Cov(X, Y) = E(XY) - E(X)E(Y)$.
$ii)$ On a $V(X + Y) = V(X) + V(Y) + 2Cov(X, Y)$. En particulier si X et Y sont indépendantes, on a $V(X + Y) = V(X) + V(Y)$.

Remarque 34 : On dispose d'inégalité de concentration classique tel que Markov, Bienaimé-Tchebytchev ou encore Jensen. On les abordera pas en détaille ici comme ce n'est pas spécifiques aux variables aléatoires discrètes.

2.2 Fonctions génératrices

Définition 35 : Soit $X : \Omega \to \mathbb{N}$ une variable aléatoire. On appelle série génératrice de X la série notée $G_X(z)$ et définie par $G_X(z) = \sum_{n=0}^{+\infty} \mathbb{P}(X = n)z^n$, de rayon de convergence ≥ 1.

Proposition 36 : Soit $X : \Omega \to \mathbb{N}$ une variable aléatoire. La série génératrice de X converge normalement pour $|z| \leq 1$, elle vérifie les propriétés suivantes :

$i)$ $z \to G_X(z)$ est définie et continue sur $|z| \leq 1$, en particulier sur $[0, 1]$.

$ii)$ Lorsque $t \in [0, 1]$, on a $G_X(t) = E(t^X)$. On a $G_X(1) = 1$.

$iii)$ X admet une espérance si et seulement si G_X est de classe C^1 sur $[0, 1]$, et dans ce cas $E(X) = G'_X(1)$.

$iv)$ X admet une variance si et seulement si G_X est de classe C^2 sur $[0, 1]$, et dans ce cas $V(X) = G''_X(1) + G'_X(1) - G'_X(1)^2$.

$v)$ Soit $Y : \Omega \to \mathbb{N}$ une variable aléatoire. Alors $G_X = G_Y$ sur $[0, 1]$ si et seulement si X et Y ont même loi.

Soit $Y : \Omega \to \mathbb{N}$ une variable aléatoire indépendante de X. Alors $G_{X+Y} = G_X G_Y$.

Développement 37 : Soit $(X_{n,i})_{(n,i)\in\mathbb{N}\times\mathbb{N}^*}$ une famille de variables aléatoires à valeur dans $\mathbb{N}$ indépendantes identiquement distribuées de loi $p_k = \mathbb{P}(X = k)$ pour $k \in \mathbb{N}$ avec $p_0 \in]0, 1[$ et d'espérance m. On définit la suite (Z_n) de la manière suivante

$$Z_0 = 1 \qquad et \qquad \forall n \in \mathbb{N},\ Z_{n+1} = \sum_{i=1}^{Z_n} X_{n,i} \quad (Z_{n+1} = 0 \quad \text{si} \quad Z_n = 0).$$

Enfin on note $\pi_n = \mathbb{P}(Z_n = 0)$ et $P_{ext} = \mathbb{P}(\exists n \in \mathbb{N}, Z_n = 0)$.
Si $m \leq 1$, $P_{ext} = 1$ et le processus s'éteint presque surement.
Si $m > 1$, $P_{ext} < 1$ et il y a une probabilité de survie non nul.

Dev 1

Remarque 38 : Ici Z_n modélise le nombre d'individus à la génération n et $X_{n,i}$ le nombre de descendant de l'individu i à la n-ième génération, π_n la probabilité d'extinction à la génération n et P_{ext} la probabilité d'extinction de la population.
On étudie alors la descendance d'un seul individu et donc on pose $Z_0 = 1$.
Le développement nous dit alors que si la moyenne de descendant pour chaque individus est inférieur ou égale à 1, alors la lignée va s'éteindre presque surement, et si la moyenne est plus grande que 1, alors il y a une probabilité non nul que la lignée survive.

3 Vers les théorèmes limites

A rajouter loi faible/forte des grands nombre. Théorème centrale limite, développement hoffding et exemples. Faire l'annexe sur papier du tableau des lois usuelles + graphe pour galton watson voir 253.

Références

1. Algèbre Gourdon
2. Probabilité 1 et 2 ouvrard
3. Probabilité Barbe