264

Variables aléatoires discrètes. Exemples et applications

I - Variables aléatoires discrètes, exemples

1) Variables aléatoires discrètes

1- Définition : Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé, et soit E un ensemble. Une application $X : \Omega \to E$ est une variable aléatoire discrète si $X(\Omega)$ est dénombrable et si pour tout $x \in E$, on a $X^{-1}(\{x\}) \in \mathcal{A}$.

2- Définition : Soit X une variable aléatoire discrète, on définit la loi de X comme une mesure de probabilité $\mathbb{P}_X$ sur $(E, \mathcal{P}(E))$ définit pour $B \in \mathcal{P}(E)$ par $\mathbb{P}_X(B) = \mathbb{P}(X \in B)$.

3- Proposition : Soit X une variable aléatoire discrète alors sa loi est caractérisée par sa valeur sur les atomes i.e. $(\mathbb{P}(X=x))_{x \in X(\Omega)}$.

4- Remarque : Soient X_1, X_2 deux variables aléatoires discrète à valeurs dans $\mathbb{R}$ alors pour $\lambda, \mu \in \mathbb{R}$, $\lambda X_1 + \mu X_2$ et $X_1 X_2$ sont des variables aléatoires discrètes.

2) Lois usuelles et leurs modélisation

Soit X une variable aléatoire (v.a)

5- Définition (Loi uniforme) Supposons $X(\Omega) = \{1, .., n\}$ $(n \geq 1)$, on dit que X est de loi uniforme si pour $k \in [\![1, n]\!]$, $\mathbb{P}(X=k) = \frac{1}{n}$. On la note $\mathcal{U}_{[\![1,n]\!]}$.

6- Exemple : Pour un lancé de dé non pipé, la variable aléatoire X égale au chiffre obtenu suit une loi uniforme sur $[\![1,6]\!]$.

7- Définition (Loi de Bernoulli) Une v.a. X à valeurs dans $\{0,1\}$ suit une loi de Bernoulli de paramètre $p \in [0,1]$ si $\mathbb{P}(X=1) = p$ et $\mathbb{P}(X=0) = 1-p$. On la note $\mathcal{B}(p)$.

8- Exemple : Le résultat d'un pile ou face avec une pièce équilibrée suit une loi de Bernoulli de paramètre $p = \frac{1}{2}$.

9- Définition (Loi Binomiale) Une v.a X à valeurs dans $\mathbb{N}$ suit une loi Binomiale, de paramètres $(n,p) \in \mathbb{N}^* \times [0,1]$, notée $\mathcal{B}(n,p)$, si pour tout $0 \leq k \leq n$, $\mathbb{P}(X=k) = \binom{n}{k} p^k (1-p)^{n-k}$

10- Exemple : Pour une urne contenant m boules (ℓ blanches, $m-\ell$ noires) alors si l'on tire au hasard avec remises n boules, le nombre de boules blanches tirées suit une loi $\mathcal{B}(n, \frac{\ell}{m})$.

11- Définition : (Loi géométrique) Une v.a X à valeurs dans $\mathbb{N}^*$ suit une loi géométrique de paramètre $p \in [0,1]$, notée $\mathcal{G}(p)$, si : $\forall k \geq 1$, $\mathbb{P}(X=k) = p(1-p)^{k-1}$

12- Exemple : Pour une épreuve aléatoire à deux issues avec une probabilité de succès p. Si on note X le temps d'attente du premier succès alors X suit une loi géométrique de paramètre p.

13- Définition : (Loi de Poisson) Une v.a X à valeurs dans $\mathbb{N}$ suit une loi de Poisson de paramètre $\lambda > 0$, notée $\mathcal{P}(\lambda)$, si $\mathbb{P}(X=k) = \frac{\lambda^k}{k!} e^{-k}$ $(\forall k \geq 0)$.

14- Exemple : C'est la loi décrivant l'apparition d'évènements rares, nombres d'accidents, nombre d'ouragan à Rennes.

3) Indépendances et sommes de variables aléatoires

15- Définition Deux variables aléatoires discrètes X_1, X_2 sont indépendantes, noté $X_1 \perp\!\!\!\perp X_2$, si et seulement si pour tout $(x_1, x_2) \in X_1(\Omega) \times X_2(\Omega)$, $\mathbb{P}(\{X_1 = x_1\} \cap \{X_2 = x_2\}) = \mathbb{P}(X_1 = x_1)\mathbb{P}(X_2 = x_2)$.

16- Remarque : Cette notion se généralise à une famille quelconque de variables aléatoires.

17- Proposition : Si X et Y sont deux variables aléatoires discrètes à valeur dans $\mathbb{Z}$ indépendantes, alors la loi de $X_1 + X_2$ est donnée par : $\forall x \in \mathbb{Z}$, $\mathbb{P}_{X+Y}(\{x\}) = \sum_{y \in \mathbb{Z}} \mathbb{P}(X=x)\mathbb{P}(Y=y-x)$

18- Exemple : Si $X \sim \mathcal{P}(\lambda_1)$, $Y \sim \mathcal{P}(\lambda_2)$ et $X \perp\!\!\!\perp Y$ alors $X + Y \sim \mathcal{P}(\lambda_1 + \lambda_2)$.
Si $X \sim \mathcal{B}(n_1, p)$, $Y \sim \mathcal{B}(n_2, p)$ et $X \perp\!\!\!\perp Y$ alors $X + Y \sim \mathcal{B}(n_1 + n_2, p)$.

19- Remarque : Si $(X_1, \dots, X_n)$ sont des v.a. de Bernoulli de paramètre p indépendantes, alors $X_1 + \dots + X_n$ est de loi $\mathcal{B}(n,p)$

20- Définition : Pour $x \in [0,1[$, on définit $D_n(x)$ le n-ième coefficient du développement dyadique de x i.e $x = \sum_{n \geq 1} \frac{D_n(x)}{2^n}$

21- Théorème : Sur l'espace probabilisé $([0,1[, \mathcal{B}_{[0,1[}, \lambda)$, les fonctions D_n définissent des variables aléatoires discrètes indépendantes, de loi de Bernoulli de paramètre $\frac{1}{2}$.
Ainsi, $\sum_{n \geq 1} \frac{D_n}{2^n}$ est une variable aléatoire de loi uniforme sur $[0,1[$.

II - Outils pour l'étude des variables aléatoires discrètes

1) Espérance et moments

22- Définition Soit X une v.a discrète à valeurs dans $\mathbb{R}$. Si la somme $\sum_{x \in \mathbb{R}} |x| \mathbb{P}(X=x)$ est finie, on dit que X est d'espérance finie et on définit alors l'espérance de X par :
$$\mathbb{E}[X] = \sum_{x \in \mathbb{R}} x \mathbb{P}(X=x).$$

23- Exemple : $\mathbb{E}[\mathcal{U}_{\{x_1, .., x_n\}}] = \frac{x_1 + .. + x_n}{n}$, $\mathbb{E}[\mathcal{B}(n,p)] = np$, $\mathbb{E}[\mathcal{G}(p)] = \frac{1}{p}$, $\mathbb{E}[\mathcal{P}(\lambda)] = \lambda$.

24- Remarque : L'espérance ne dépend que de la loi de X.

25- Théorème : Soit X une v.a discrète à valeurs dans E et $f : E \to \mathbb{R}$. Alors $Y = f \circ X$ est une variable aléatoire réelle discrète. De plus Y admet une espérance si et seulement si $\sum_{x \in E} |f(x)| \mathbb{P}(X=x)$

Alors l'espérance est donnée par : $\mathbb{E}[f(X)] = \sum_{x \in E} f(x) \mathbb{P}(X=x)$.

26 - Proposition (Inégalité de Markov) Soit X une v.a. réelle d'espérance finie. Alors $\forall \varepsilon>0, \mathbb{P}(|X|>\varepsilon) \leq \frac{\mathbb{E}[X]}{\varepsilon}$.

27 - Définition Pour $p \geq 1$, une v.a discrète réelle X admet un moment d'ordre p ssi X^p est d'espérance finie. Le moment d'ordre p de X est alors $\mathbb{E}[X^p]$.

28 - Proposition : Soit X une v.a. discrète possédant un moment d'ordre 2. Alors X est d'espérance finie et la v.a. $X-\mathbb{E}[X]$ possède un moment d'ordre 2. On définit alors la variance de X comme $var(X) = \mathbb{E}[(X-\mathbb{E}[X])^2]$ et on a : $var(X) = \mathbb{E}[X^2]-\mathbb{E}[X]^2$

29 - Remarque : Avec le théorème de transfert on a : $var(X) = \sum_{x\in\mathbb{R}} x^2\,\mathbb{P}(X=x) - \mathbb{E}[X]^2$

30 - Exemple : $Var(\mathcal{U}_{\{x_1,\dots,x_n\}}) = \frac{x_1^2+\dots+x_n^2}{n} - \left(\frac{x_1+\dots+x_n}{n}\right)^2$, $Var(\mathcal{B}(n,p)) = np(1-p)$, $Var(\mathcal{G}(p)) = \frac{1-p}{p^2}$, $Var(\mathcal{P}(\lambda)) = \lambda$.

31 - Proposition (Inégalité de Tchebychev) Soit X une v.a réelle discrète de variance finie. Alors $\forall \varepsilon>0, \mathbb{P}(|X-\mathbb{E}[X]| \leq \varepsilon) \leq \frac{Var(X)}{\varepsilon^2}$.

32 - Application (Théorème de Weierstrass) Pour $f:[0,1] \longrightarrow \mathbb{R}$ continue on définit pour $n\in\mathbb{N}$, le n-ième polynôme de Bernstein par $B_n(x) = \sum_{k=0}^{n} \binom{n}{k} x^k(1-x)^{n-k} f\left(\frac{k}{n}\right)$.
Alors B_n converge uniformément vers f sur $[0,1]$ et il existe $C>0$ telle que $\|f-B_n\|_\infty \leq C\,\omega\left(\frac{1}{\sqrt{n}}\right)$ $(\forall n\geq 1)$ où pour $h>0$, $\omega(h) = \sup\{|f(x)-f(y)|, |x-y|\leq h\}$ est le module de continuité de f.

2) Fonction génératrice

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

33 - Définition / Proposition Pour tout $s\in]-1,1[$, la variable aléatoire s^X est bien définie et d'espérance finie. On note quand cette quantité existe, $G_X(s) = \mathbb{E}[s^X]$, la fonction génératrice de X.

34 - Proposition : La fonction génératrice vérifie :

i) Le domaine de définition de G_X contient l'intervalle $[-1,1]$, et on a :
$$\forall s\in[-1,1], |G_X(s)|\leq 1 \text{ et } G_X(1)=1$$
(ii) Pour tout $s\in[-1,1]$, on a :
$$G_X(s) = \sum_{n=0}^{+\infty} s^n\,\mathbb{P}(X=n)$$
(iii) G_X est continue sur $[-1,1]$ et de classe $\mathcal{C}^\infty$ sur $]-1,1[$.

(iv) La restriction de G_X à un voisinage de 0 caractérise la loi de X, on a, pour tout $n\in\mathbb{N}$:
$$\mathbb{P}(X=n) = \frac{G_X^{(n)}(0)}{n!}$$

35 - Exemple : Si $X\sim\mathcal{B}(n,p)$ alors $\forall s\in\mathbb{R}, G_X(s) = (ps+1-p)^n$.
Si $X\sim\mathcal{P}(\lambda)$ alors $\forall s\in\mathbb{R}, G_X(s) = \exp(\lambda(s-1))$.
Si $X\sim\mathcal{G}(p)$ alors $\forall s\in[-1,1], G_X(s) = \frac{ps}{1-(1-p)s}$.

36 - Proposition : Soient X et Y deux variables aléatoires indépendantes à valeurs dans $\mathbb{N}$. Alors la fonction génératrice de $X+Y$ est donnée pour $s\in[-1,1]$ par :
$$G_{X+Y}(s) = G_X(s)\,G_Y(s).$$

37 - Application : (Galton-Watson) Soient $X_{n,k}$ des v.a. iid de loi discrète sur $\mathbb{N}$, ayant une espérance finie. On définit la suite de v.a. $(Z_n)_{n\in\mathbb{N}}$ par $Z_0=1$ et $Z_{n+1} = \sum_{k=1}^{Z_n} X_{n+1,k}$.
Alors la probabilité d'extinction $a = \mathbb{P}(\{\exists n\in\mathbb{N} \text{ tq } Z_n=0\})$ vaut 1 si $\mathbb{E}[X_{1,1}]<1$ ou si $\mathbb{P}(X_{1,1}=1)=1$ et est dans $[0,1[$ si $\mathbb{E}[X_{1,1}]\geq 1$ et $\mathbb{P}(X_{1,1}=1)<1$.

III - Comportement limite

38 - Définition : Soient $(X_n)_{n\in\mathbb{N}}$ et X des v.a. réelles. On dit que $(X_n)_{n\in\mathbb{N}}$ converge en loi vers X si et seulement si pour toute fonction continue bornée sur $\mathbb{R}$, on a : $\mathbb{E}[f(X_n)] \xrightarrow[n\to+\infty]{} \mathbb{E}[f(X)]$. On note $X_n \xrightarrow{\mathcal{L}} X$.

39 - Proposition : Si $(X_n)_{n\in\mathbb{N}}$ et X sont à valeurs dans $\mathbb{N}$, alors la suite $(X_n)_n$ converge en loi vers X si et seulement si $\forall k\geq 0, \mathbb{P}(X_n=k) \xrightarrow[n\to+\infty]{} \mathbb{P}(X_n=k)$.

40 - Remarque : Cette propriété reste vraie si les $(X_n)_{n\in\mathbb{N}}$ et X sont à valeurs dans un espace dénombrable discret.

41 - Théorème : Soit $(X_n)_{n\in\mathbb{N}}$ une suite à valeurs dans $\mathbb{N}$, de fonctions génératrices G_{X_n}. Et si X est une variable aléatoire à valeurs dans $\mathbb{N}$ de fonction génératrice G_X. Alors on a équivalence entre :

1) $X_n \xrightarrow{\mathcal{L}} X$

2) $G_{X_n} \longrightarrow G_X$ simplement sur $[0,1[$.

42 - Théorème : Convergence en loi de la loi binomiale.

Soit $(X_n)_{n\in\mathbb{N}}$ une suite de variable aléatoires tel que $X_n \sim \mathcal{B}(n,p_n)$. On suppose que $np_n \xrightarrow[n\to\infty]{} \lambda$.

Alors X_n converge en loi vers la loi de Poisson de paramètre λ.

43 - Remarque : Le théorème nous dit que si n "est grand" et np_n "pas trop grand", on peut remplacer la loi binomiale par une loi de Poisson. Cela justifie pourquoi la loi de Poisson est une bonne modélisation pour les évènements rares.

44 - Théorème : Convergence en loi de la loi hypergéométrique.

Soit $(X_j)_{j\in\mathbb{N}}$ une suite de variable aléatoires tel que X_j suit une loi hypergéométrique de paramètre (N_j, m_j, k). On suppose que $N_j \to +\infty$ et $\frac{m_j}{N_j} \xrightarrow[j\to+\infty]{} p$. Alors (X_j) converge en loi vers une variable de loi binomiale $\mathcal{B}(k,p)$.

45 - Théorème : (Théorème central limite)

Soient $(X_n)_{n\in\mathbb{N}}$ des variables aléatoires réelles indépendantes et de même loi, possédant un moment d'ordre 2. Alors si on note $S_n = X_1 + \dots + X_n$ on a :

$$\frac{S_n - n\,\mathbb{E}[X_1]}{\sqrt{n\mathbb{V}ar(X_1)}} \xrightarrow[n\to+\infty]{} \mathcal{N}(0,1)$$

47 - Application : Pour S_n de loi $\mathcal{B}(n,p)$ pour tout $a<b\in\mathbb{R}$, on a :

$$\mathbb{P}(a<S_n<b) - \frac{1}{\sqrt{2\pi}}\int_{\frac{a-np}{\sqrt{np(1-p)}}}^{\frac{b-np}{\sqrt{np(1-p)}}} e^{-\frac{x^2}{2}}dx \xrightarrow[n\to\infty]{} 0$$

48 - Théorème (Loi forte des grands nombres) Soient (X_n) des v.a. réelles indépendantes et de même loi, admettant un moment d'ordre 2. Alors la suite de v.a $S_n = \frac{X_1+\dots+X_n}{n}$ converge $\mathbb{P}$-presque partout vers $\mathbb{E}[X_1]$.

49 - Application : La moyenne empirique $S_n = \frac{X_1+\dots+X_n}{n}$ permet ainsi d'estimer $\mathbb{E}[X_1]$.

IV - Utilisation en algèbre et théorie des nombres

1) Etudes des permutations aléatoires

Soit X_n une variable aléatoire à valeur dans $\mathfrak{S}_n$ de loi uniforme.

50 - Proposition : Espérance et variance du nombre de point fixes

Soit Y_n la variable aléatoire comptant le nombre de point fixes de X_n. Alors on a que : $\mathbb{E}[Y_n] = \mathbb{V}ar[Y_n] = 1$

51 - Proposition : Probabilités d'être un dérangements.

Soit $\mathcal{D}_n$ l'ensemble des dérangements de $\mathfrak{S}_n$. Alors

$$\mathbb{P}(X_n \in \mathcal{D}_n) = 1 - \sum_{p=1}^{n} \frac{(-1)^{p+1}}{p!} \xrightarrow[p\to\infty]{} \frac{1}{e}$$

52 - Proposition : Décomposition en produit de cycle.

Pour tout $j\in[\![1,n]\!]$ on note $C_j^{(n)}$ la variable aléatoire comptant le nombre de cycles de taille j dans la décomposition en cycles disjoints de X_n.

__Les lois conjointes__ :

$$\mathbb{P}(C_1^{(n)}=c_1,\dots,C_n^{(n)}=c_n) = \mathbb{1}_{\{\sum_{j=1}^n jc_j = n\}} \prod_{j=1}^{n} \left(\frac{1}{j}\right)^{c_j}\frac{1}{c_j!}$$

__Lois marginales__ : Pour $1\leq j\leq n$,

$$\mathbb{P}(C_j^{(n)}=k) = \frac{j^{-k}}{k!}\sum_{\ell=0}^{\lfloor \frac{n}{j}\rfloor - k} (-1)^\ell \frac{j^{-\ell}}{\ell!}$$

__Lois limites__ : $C_j^{(n)}$ converge en loi vers une variable aléatoires de loi de Poisson de paramètre $\frac{1}{j}$.

2) Théorie des nombres.

Pour $k\in\mathbb{N}$ on note $\upsilon(k)$ le nombre de diviseurs premiers de k. De plus pour $n\in\mathbb{N}$ on considère X_n une variable aléatoire à valeurs dans $[\![1,n]\!]$ de loi uniforme.

53 - Proposition : Soit $\varphi:\mathbb{N}\to\mathbb{R}_+$ tel que $\varphi(n)\xrightarrow[n\to\infty]{} 0$ alors on a :

$$\mathbb{P}\left(|\upsilon(X_n) - \log(\log(n))| > \varphi(n)\sqrt{\log(\log(n))}\right) \xrightarrow[n\to+\infty]{} 0$$

54 - Théorème : Erdös-Kac (admis) On a

$$\frac{\upsilon(X_n) - \log(\log(n))}{\sqrt{\log(\log(n))}} \xrightarrow{\mathcal{L}} \mathcal{N}(0,1)$$

Annexes :

Lois	$\mathcal{B}(n,p)$	$\mathcal{P}(\lambda)$	$\mathcal{G}(p)$
Espérance	np	λ	$\frac{1}{p}$
Variance	$np(1-p)$	λ	$\frac{1-p}{p^2}$
Fonction génératrice	$(ps+(1-p))^n$	$e^{\lambda(s-1)}$	$\frac{ps}{1-(1-p)s}$

Références :

Ouvrard, Probabilités 1 & 2

Garet, Kurtzmann, De l'intégration aux probabilités

Appel, Probabilités pour les non probabilistes

Terence Tao, Van H. Vu, Additive Combinatorics (IV.2 Théorie des nombres)

Arratia, Barbour, Tavaré, Logarithmic combinatorial structures: a probabilistic approach

Chafaï, Malrieu, Recueil de Modèles aléatoires

(IV.1 Permutations aléatoires)