Leçon 264 Variables aléatoires discrètes, exemples et applications

Dorian Cacitti-Holland

2020-2021

Références.

1. Probabilités tome 1 et 2 de Jean-Yves Ouvrard
2. Probabilité de Barbe et Ledoux
3. Analyse de Queffélec et Zuily
4. Probabilités et statistiques de Chabanol et Ruch
5. Les contre-exemples en mathématiques de Bertrand Hauchecorne

Développements.

1. Théorème de Weierstrass par les polynômes de Bernstein
2. Marche aléatoire sur $\mathbb{Z}^d$

Table des matières

1 Variables aléatoires discrètes

1.1 Lois discrètes

(Chapitres 8.2 de Probabilités tome 2 de Jean-Yves Ouvrard et II.3 et III.1 de Probabilité de Barbe et Ledoux)

On considère $(\Omega, \mathcal{A}, \mathbb{P})$ espace probabilisé et X variable aléatoire de $(\Omega, \mathcal{A}, \mathbb{P})$ dans $(\mathbb{N}, \mathcal{P}(\mathbb{N}))$.

1. Définition : La loi de X est la mesure de probabilité $\mathbb{P}_X$ définie par $\forall B \in \mathcal{B}, \mathbb{P}_X(B) = \mathbb{P}(\{\omega \in \Omega, X(\omega) \in B\}) = \mathbb{P}(X \in B)$
2. Exemple : δ_x est la mesure de probabilité de la variable aléatoire constante x
3. Définition : On dit que la loi de X est discrète si $\mathbb{P}_X$ est une combinaison linéaire finie ou dénombrable de masse de Dirac, ie $\mathbb{P}_X = \sum\limits_{i \in I} p_i \delta_{x_i}$ avec $p_i = \mathbb{P}(X = x_i)$
4. Remarque : Dans ce cas X est à valeurs dans un espace fini ou dans un espace dénombrable ie en bijection avec $\mathbb{N}$

1.2 Lois usuelles

(Appendice de Probabilité de Barbe et Ledoux)

1. Définition : On dit que X suit une loi de Bernoulli $\mathcal{B}(p)$ si $\mathbb{P}_X = p\delta_1 + (1-p)\delta_0$
2. Remarque : Dans ce cas X représente le résultat pouvant valoir 0 ou 1 uniquement
3. Exemple : Le lancer d'une pièce
4. Définition : On dit que X suit une loi binomiale $\mathcal{B}(n, p)$ si $\mathbb{P}_X = \sum\limits_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} \delta_k$
5. Remarque : Dans ce cas X représente le nombre de succès dans la répétition de n expériences indépendantes identiquement distribuées de loi $\mathcal{B}(p)$
6. Exemple : On compte le nombre de faces obtenues après n lancers de pièces
7. Définition : On dit que X suit une loi géométrique $\mathcal{G}(p)$ si $\mathbb{P}_X = \sum\limits_{k=1}^{+\infty} (1-p)^k p \delta_k$
8. Remarque : Dans ce cas X représente le rang du premier succès d'une suite de variables aléatoires indépendantes identiquement distribuées de loi $\mathcal{B}(p)$
9. Exemple : La première fois qu'on a face dans une série de lancers de pièces
10. Définition : On dit que X suit une loi de Poisson $\mathcal{P}(\lambda)$ si $\mathbb{P}_X = \sum\limits_{k=0}^{+\infty} e^{-\lambda} \frac{\lambda^k}{k!} \delta_k$
11. Remarque : Dans ce cas X représente le comportement de la limite d'une suite de variables aléatoires indépendantes identiquement distribuées de loi $\mathcal{B}\left(n, \frac{\lambda}{n}\right)$

2 Propriétés particulières dans le cas discret

2.1 Indépendance

(Chapitres IV de Probabilité de Barbe et Ledoux, 9.1 de Probabilités tome 2 de Jean-Yves Ouvrard et XIII.II.1.c d'Analyse de Queffélec et Zuily)

On considère $(X_i)_{i\in I}$ une famille de variables aléatoires discrètes.

1. Définition : On dit que $(X_i)_{i\in I}$ est indépendante si pour tout $J \subset I$ fini et tout $B_j \in \mathcal{P}(\mathbb{N})$, $\mathbb{P}\left(\bigcap_{j\in J}(X_j \in B_j)\right) = \prod_{j\in J}\mathbb{P}(X_j \in B_j)$
2. Exemple : On considère $A_n = \bigcup_{1\leq k\leq 2^{n-1}} \left]\frac{2(k-1)}{2^n}, \frac{2k-1}{2^n}\right]$ et $X_n = \mathbb{1}_{A_n}$, alors les X_n sont indépendants et de même loi $\mathcal{B}\left(\frac{1}{2}\right)$
3. Théorème : Soit X_1, X_2 discret, alors (X_1, X_2) est discret et X_1, X_2 sont indépendants si et seulement si $\forall x_1 \in X_1(\Omega), x_2 \in X_2(\Omega), \mathbb{P}(X_1 = x_1, X_2 = x_2) = \mathbb{P}(X_1 = x_1)\mathbb{P}(X_2 = x_2)$
4. Remarque : L'indépendance deux à deux de X_1, X_2, X_3 ne suffit pas à montrer que X_1, X_2, X_3 sont indépendants
5. Exemple : Soit X, Y les résultats de lancers indépendants d'un dé à six faces et $Z = \mathbb{1}_{X+Y\in 2\mathbb{N}}$, alors X, Y, Z sont deux à deux indépendants mais pas indépendants
6. Exemple : Si X, Y indépendantes de lois $\mathcal{P}(\lambda)$ et $\mathcal{P}(\mu)$ alors $X + Y \sim \mathcal{P}(\lambda + \mu)$
7. Exemple : Si $X_1, ..., X_n \sim \mathcal{B}(p)$ indépedantes alors $S_n = \sum_{k=1}^{n} X_k \sim \mathcal{B}(n, p)$
8. Définition : Soit $f \in C([0, 1)$, alors les polynômes de Bernstein de f sont $B_n(f) = \sum_{k=0}^{n}\binom{n}{k}X^k(1-X)^{n-k}f\left(\frac{k}{n}\right) = \mathbb{E}\left(f\left(\frac{S_n}{n}\right)\right)$
9. Lemme : Inégalité de Khintchine : Soit $Y_1, ..., Y_n$ n variables aléatoires indépendantes sur $(\Omega, \mathcal{A}, \mathbb{P})$ de loi uniforme sur $\left\{-\frac{1}{2}, \frac{1}{2}\right\}$ (des variables de Rademacher) et $Y \in F_n$ le $\mathbb{R}$-espace vectoriel de dimension n engendré par les $Y_1, ..., Y_n$, alors $\sqrt{\mathbb{E}(Y^2)} \leq \sqrt{e}\mathbb{E}(|Y|)$
10. Application ; Théorème de Weierstrass : Soit $[a, b] \subset \mathbb{R}$, alors les fonctions polynomiales restreintes à $[a, b]$ sont denses dans $C([a, b], \mathbb{R})$ pour la norme infini

2.2 Moments

(Chapitres III.2 et III.3 de Probabilités et statistiques de Chabanol et Ruch)
On considère X à valeurs dans $\mathbb{N}$.

1. Définition : L'espérance de X est $\mathbb{E}(X) = \int_\Omega X(\omega)d\mathbb{P}_X(\omega)$
2. Proposition : $\mathbb{E}(X) = \sum_{k=0}^{+\infty} n\mathbb{P}(X = n)$
3. Exemple : Soit $B \in \mathcal{A}$ et $X = \mathbb{1}_B$, alors $\mathbb{E}(X) = \mathbb{P}(B)$
4. Exemple : Les espérances des lois usuelles sont $p, np, \frac{1}{p}, \lambda$
5. Remarque : Il existe des variables aléatoires discrètes n'admettant pas d'espérance
6. Exemple : Si $\mathbb{P}_X = \sum_{n=1}^{+\infty} \frac{1}{2^n}\delta_{2^n}$ alors X n'admet pas d'espérance (Exemple 19.3 des Contre-exemples de Bertrand Hauchecorne)
7. Théorème : X admet un moment d'ordre k si et seulement si $\sum |n|^k\mathbb{P}(X = n)$ converge, dans ce cas $\mathbb{E}(X^k) = \sum_{n=0}^{+\infty} n^k\mathbb{P}(X = n)$

2.3 Fonctions génératrices

(Chapitre 5.3 de Probabilités tome 1 de Jean-Yves Ouvrard et IV.7 de Probabilités et statistiques de Chabanol et Ruch)

On considère X variable aléatoire à valeurs dans $\mathbb{N}$.

1. Définition : $G_X = \sum_{n=0}^{+\infty} \mathbb{P}(X = n) s^n$
2. Exemple : Si $X \sim \mathcal{B}(n,p)$ alors $G_X = (sp + 1 - p)^n$
3. Théorème : G_X est continue sur $[-1,1]$ et de classe C^∞ sur $]-1,1[$ et $\forall n \in \mathbb{N}, \mathbb{P}(X = n) = \frac{G^{(n)}(0)}{n!}$
4. Corollaire : G_X caractérise la loi de X
5. Proposition : Si X, Y indépendantes alors $G_{X+Y}(t) = G_X(t)G_Y(t)$
6. Théorème : X admet un moment d'ordre k si et seulement si G_X est k-fois dérivable à gauche en 1, dans ce cas pour $k = 1$, $\mathbb{E}(X) = G'_X(1)$
7. Application : Soit $(X_n)_{n\in\mathbb{N}}$ suite de variables aléatoires discrètes, T variable aléatoire dans $\mathbb{N}$ toutes indépendantes, $S_n = \sum_{k=1}^{n} X_k$ et $S = S_T$, alors $G_S = G_T \circ G_{X_1}$
8. Théorème : $(G_{X_n})_{n\in\mathbb{N}}$ converge simplement vers G_X si et seulement si $X_n \xrightarrow[n\to+\infty]{loi} X$
9. Application : Si $X_n \sim \mathcal{B}\left(n, \frac{\lambda}{n}\right)$ alors $X_n \xrightarrow[n\to+\infty]{loi} \mathcal{P}(\lambda)$
10. Théorème de Moivre-Laplace : Si $X_n \sim \mathcal{B}(n,p)$, alors $\frac{X_n - np}{\sqrt{np(1-p)}} \xrightarrow[n\to+\infty]{loi} \mathcal{N}(0,1)$

3 Cas des chaînes de Markov

3.1 Définitions et exemples

(Chapitres VII.1 et VII.2 de Probabilités et statistiques de Chabanol et Ruch)

On considère $X = (X_n)_{n\in\mathbb{N}}$ suite de variables aléatoires dans E au plus dénombrable.

1. Définition : On dit que X est une chaîne de Markov (homogène) si $\mathbb{P}(X_{k+1} = x_{k+1} \mid X_k = x_k, ..., X_0 = x_0) = \mathbb{P}(X_{k+1} = x_{k+1} \mid X_k = x_k)(= \mathbb{P}(X_1 = x_{k+1} \mid X_0 = x_k) = p_{x_k,x_{k+1}})$, avec $p_{x,y}$ probabilité de transition pour aller de x à y
2. Lemme : Soit ν_0 la loi de X_0, alors $\mathbb{P}(X_n = x_n, ..., X_0 = x_0) = \nu_0(x_0) \prod_{k=0}^{n-1} p_{x_k,x_{k+1}}$
3. Définition : $P = (p_{x,y})_{x,y\in E}$ est appelée matrice de transition de X
4. Proposition : La loi de X est caractérisée par ν_0 et P, plus précisément $\nu_n = \nu_0 P^n$
5. Exemple : Chaîne à deux états : Si $E = \{1,2\}$ alors $P = \begin{pmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{pmatrix}$ avec $0 < \alpha, \beta < 1$, ainsi le graphe associé en annexe est simple
6. Exemple : Marche aléatoire sur $\mathbb{Z}$: Si $E = \mathbb{Z}$ et $p_{i,i-1} = p$, $p_{i,i+1} = p$ alors le graphe associé est en annexe

3.2 Propriétés et caractéristiques

(Chapitres VII.3, VII.4, VII.5 et VII.6 de Probabilités et statistiques de Chabanol et Ruch)

1. Théorème : Relation de Chapman-Kolmgorov : $P^{m+n} = P^m P^n$ ie $\mathbb{P}(X_{n+m} = j \mid X_0 = i) = \sum\limits_{k \in E} \mathbb{P}(X_m = k \mid X_0 = i)\mathbb{P}(X_n = j \mid X_0 = k)$
2. Définition : On dit que j est accessible à partir de i si $p_{i,j}^{(n)} > 0$ et on dit que i et j communiquent l'accessibilité est dans les deux sens
3. Proposition : La relation de communication est une relation d'équivalence, les classes d'équivalences sont appelés les composantes irréductibles de la chaine
4. Définition : On dit que la chaîne est irréductible si tous les états communiquent entre eux, autrement dit il n'existe qu'une seule composante irréductible
5. Exemple : La marche aléatoire sur $\mathbb{Z}$ est irréductible
6. Définition : La période de j est $d(j) = pgcd(n, p_{j,j}^{(n)} > 0)$, si $d(j) = 1$ alors on dit que j est apériodique et on dit que la chaîne est apériodique si tout état est apériodique
7. Proposition : La périodicité est une propriété de classe
8. Exemple : La chaîne d'Ehrenfest en annexe est de période 2
9. Définition : $T_j := inf(n \geq 1, X_n = j), N_j = \sum\limits_{n=0}^{+\infty} \mathbb{1}_{(X_n = j)}$ et on dit que j est récurrent si $\mathbb{P}_j(T_j < +\infty) = 1$, sinon on dit que j est transitoire
10. Proposition : Si j est récurrent alors $\mathbb{P}_j(N_j = +\infty) = 1$ donc $\mathbb{E}_j(N_j) = +\infty$, et si j est transitoire alors $\mathcal{L}(N_j \mid X_0 = j) = \mathcal{G}(p)$
11. Théorème : j récurrent si et seulement si $\sum p_{j,j}^{(n)}$ diverge
12. Théorème : La marche alétoire sur $\mathbb{Z}^d$ est récurrente si et seulement si $d \leq 2$

3.3 Probabilités invariantes et comportement asymptotique

(Chapitre VII.7 de Probabilités et statistiques de Chabanol et Ruch)

1. Définition : On dit que μ est une loi de probabilité stationnaire si $\nu = \nu P$
2. Proposition : Dans ce cas si j transitoire alors $\nu(j) = 0$
3. Théorème : Si E fini alors il existe une loi de probabilité stationnaire, si de plus X irréductible alors elle est unique donnée par $\nu(i) = \frac{1}{\mathbb{E}_i(T_i)}$, si de plus X est apériodique alors $P^n \underset{n \to +\infty}{\longrightarrow} {}^t(\nu, ..., \nu)$
4. Exemple : Pour la chaîne à deux états la loi de probabilité stationnaire est $\nu(1) = \frac{\beta}{\alpha+\beta}, \nu(2) = \frac{\alpha}{\alpha+\beta}$
5. Théorème : Dans ce cas, avec ν son unique loi de probabilité stationnaire, soit f fonction sur E, alors $\frac{1}{n}\sum\limits_{k=0}^{n-1} f(X_k) \underset{n \to +\infty}{\xrightarrow{PS}} \sum\limits_{i \in E} \frac{f(i)}{\mathbb{E}_i(T_i)} = \int_E f d\mu$

6. Corollaire : $\frac{1}{n}\sum_{k=0}^{n-1} \mathbb{1}_{(X_k=i)} \xrightarrow[n\to+\infty]{PS} \mu(i)$

7. Théorème : $\sqrt{n}\left(\frac{1}{n}\sum_{k=0}^{n-1} f(X_k) - \int_E f d\mu\right) \xrightarrow[n\to+\infty]{loi} \mathcal{N}(0,\sigma^2)$

4 Estimation d'une loi discrète

4.1 Construction du test du χ^2

(Chapitre XIV.1 de Probabilités et statistiques de Chabanol et Ruch)
On considère un n-échantille discret $X_1, ..., X_n$ à valeurs dans $[\![1, m]\!]$ de loi $\pi = (\pi_1, ..., \pi_m)$.

1. Définition : Le nombre de fois que la valeur i apparaît dans l'échantillon est $N_i(n) := \sum_{k=1}^{n} \mathbb{1}_{(X_k=i)}$, de plus on note $N(n)$ le vecteur aléatoire correspondant

2. Théorème : $\frac{N(n)-n\pi}{\sqrt{n}} \xrightarrow[n\to+\infty]{loi} \mathcal{N}(0, \Delta_\pi - {}^t\pi\pi)$ où $\delta_\pi = diag(\pi_1, ..., \pi_m)$

3. Corollaire : $D_n(\pi) := \sum_{i=1}^{m} \frac{(N_i(n)-n\pi_i)^2}{n\pi_i} \xrightarrow[n\to+\infty]{loi} \chi^2(m-1)$

4.2 Tests d'ajustement et d'indépendance

(Chapitres XIV.2 et XIV.5 de Probabilités et statistiques de Chabanol et Ruch)
On considère p une loi de probabilité discrète sur $[\![1, m]\!]$.

1. Proposition : Si $H_1 : \pi \neq p$ alors $\frac{N(n)}{n} \xrightarrow[n\to+\infty]{PS} \pi \neq p$, donc $D_n(p) = \sum_{i=1}^{m} \frac{n}{p_i}\left(\frac{N_i(n)}{n} - p_i\right)^2 \xrightarrow[n\to+\infty]{PS} +\infty$

2. Application : Soit $\alpha \in]0, 1[$, on considère la région de réjet $]q_\alpha, +\infty[$ avec q_α quantile de la loi $\chi^2(m-1)$, alors $D_n(p) > q_\alpha$ alors on rejette l'hypothèse $H_0 : \pi = p$ contre $H_1 : \pi \neq p$, sinon on ne rejette pas l'hypothèse

3. Exemple : On lance une pièce 200 fois et on obtient 110 piles, alors on ne rejette pas l'hypothèse $H_0 : p = \frac{1}{2}$ contre $H_1 \neq \frac{1}{2}$ avec $\alpha = 0,05$

4. Définition : On observe un n-échantillon $((X_1, Y_1), ..., (X_n, Y_n))$ à valeurs dans un ensemble $\{(x(i), y(j), 1 \leq i \leq r, 1 \leq j \leq s\}$, alors le nombre d'observations $(x(i), y(j))$ est $N_{i,j} := \sum_{k=1}^{n} \mathbb{1}_{((X_k,Y_k)=x(i),y(j))}$, de plus $N_i := \sum_{j=1}^{s} N_{ij}$ et $N^j := \sum_{i=1}^{r} N_{ij}$

5. Proposition : $n\sum_{i=1}^{r}\sum_{j=1}^{s} \frac{\left(N_{ij} - \frac{N_i N^j}{n}\right)^2}{N_i N^j} \xrightarrow[n\to+\infty]{loi} \chi^2((r-1)(s-1))$

6. Application : On peut comme précédemment construire une région de rejet de l'hypothèse $H_0 : X \perp Y$ contre $H_1 : H_0^c$