
V.A.R. discrètes finies

Leçon 4

EN2D2- Lycée Gustave Eiffel - Bordeaux
Thierry Sageaux.

Résumé

Notions de base complètes sur les variables aléatoires discrètes finies avant d'attaquer les continues...

Table des matières

I Généralités sur les variables aléatoires	1
I.1 Définition et premières propriétés	1
I.2 Couples de variables aléatoires	3
II Variables composées	4
III Lois usuelles	5
III.1 Loi uniforme : $\mathcal{U}(n)$	5
III.2 Loi de Bernoulli : $\mathcal{B}(p)$	5
III.3 Loi binomiale : $\mathcal{B}(n, p)$	5
III.4 Loi Multinomiale.	5
III.5 Loi hypergéométrique : $\mathcal{H}(N, n, p)$	6

Dans tout ce chapitre, $(\Omega, \mathcal{T})$ est un espace probabilisable.

I Généralités sur les variables aléatoires

I.1 Définition et premières propriétés

Définition 4.1 On appelle *variable aléatoire discrète* X toute application définie sur Ω à valeurs dans $\mathbb{R}$. L'ensemble image $X(\Omega) = \{x_i, i \in I\}$ est l'ensemble des valeurs prises par X . Cet ensemble est dénombrable (i.e. $I = \mathbb{N}$) ou fini et on dit dans ce cas que X est une variable aléatoire discrète *finie*.

$$\begin{array}{ccc} X : & \Omega & \longrightarrow \mathbb{R} \\ & \omega & \longmapsto x \end{array}$$

Notation : $(X = x) = X^{-1}(x) = \{\omega \in \Omega / X(\omega) = x\}$.

Définition 4.2 On appelle *loi de probabilité* de X l'application :

$$\begin{array}{ccc} p : & X(\Omega) & \longrightarrow [0, 1] \\ & x_i & \longmapsto p(X = x_i) = p_i \end{array}$$

On la définit donc via les probabilités des évènements élémentaires.

Un cas particulier de variable aléatoire est donné par la définition suivante :

Définition 4.3 Soit $(\Omega, \mathcal{T}, p)$ un espace probabilisé. Soit $A \in \mathcal{T}$ un événement. On peut alors définir la **variable aléatoire indicatrice de A** par la variable aléatoire X_A (attention, il n'existe pas de notation standardisée) qui vaut 1 si A est réalisé et 0 sinon.

Ainsi, $p(X_A = 1) = p(A)$ et $p(X_A = 0) = 1 - p(A)$.

Définition 4.4 On appelle **espérance mathématique** de X le réel :

$$E(X) = \sum_{i \in I} x_i p_i$$

Définition 4.5 On appelle **variance** de X le réel

$$V(X) = \sum_{i \in I} (x_i - E(X))^2 p_i$$

On appelle **écart type** de X le réel

$$\sigma(X) = \sqrt{V(X)}$$

En étendant les notations à la variable aléatoire X^2 , on a $E(X^2) = \sum_{i \in I} x_i^2 p_i$
et plus généralement,

Définition 4.6 Si $r \in \mathbb{N}$ et X est une variable aléatoire, on appelle **moment d'ordre r de X** , noté $m_r(X)$, l'espérance de la variable X^r . Ainsi,

$$m_r(X) = \sum_{i=1}^n x_i^r p(X = x_i) = E(X^r).$$

On utilise aussi parfois le **moment centré d'ordre r** , noté $\mu_r(X)$ qui est l'espérance mathématique de la variable aléatoire $(X - E(X))^r$:

$$\mu_r(X) = \sum_{i=1}^n (x_i - E(X))^r p(X = x_i) = E[(X - E(X))^r].$$

Remarque: Le moment d'ordre 1 est donc l'espérance et le moment centré d'ordre 2 est la variance.

Théorème 4.7 (Formule de König-Huygens)

$$V(X) = E(X^2) - (E(X))^2$$

Procédé mnémotechnique: Un moyen simple de s'en souvenir et de le dire à haute voix en français : "La moyenne des carrés moins le carré de la moyenne". C'est joli, ça sonne bien, en tout cas beaucoup mieux que le contraire : "Le carré de la moyenne moins la moyenne des carrés" dans laquelle l'allitération des trois "m" sonne mal.

Définition 4.8 On appelle **fonction de répartition** de X l'application

$$\begin{aligned} F : \mathbb{R} &\longrightarrow [0, 1] \\ x &\longmapsto p(X \leq x) \end{aligned}$$

Remarques: • F est croissante sur $\mathbb{R}$.

- $\lim_{x \rightarrow -\infty} F(x) = 0$ et $\lim_{x \rightarrow +\infty} F(x) = 1$.
- $F(x) = \sum_{x_i \leq x} p_i$
- $P(X = x_i) = F(x_i) - F(x_{i-1})$.

Exercice 1. ♪

- 1) Soit X une variable aléatoire à valeurs dans $\llbracket 1, n \rrbracket$. Montrer que $E(X) = \sum_{k=1}^n p(X \geq k)$.
- 2) On tire deux jetons simultanément dans un sac contenant 10 jetons numérotés de 1 à 10. On appelle X la variable aléatoire désignant le minimum des numéros tirés. Calculer l'espérance de X .

I.2 Couples de variables aléatoires

Exercice 2. ♪

Soit (X, Y) un couple de variables aléatoires dont la loi est définie par le tableau suivant :

Y, X	0	1
0	p	$\frac{1}{2} - p$
1	$\frac{1}{3} - p$	$\frac{1}{6} + p$

- 1) A quel intervalle doit appartenir p pour que ces données soient acceptables ?
- 2) Déterminer $E(X)$, $E(Y)$, $V(X)$, $V(Y)$.
- 3) Déterminer p pour que X et Y soient indépendantes.

Remarque: La loi de probabilité d'un couple (X, Y) s'appelle **loi conjointe de X et de Y** alors que les variables aléatoires X et Y s'appellent les **lois marginales**.

Exercice 3. ♪

Montrer que si deux variables aléatoires X et Y sont indépendantes, alors $E(XY) = E(X)E(Y)$.

Définition 4.9 On appelle **covariance** de X et Y le réel

$$Cov(X, Y) = E(XY) - E(X) \times E(Y)$$

Nota Bene : $Cov(X, X) = V(X)$.

Remarque: La covariance permet d'évaluer le sens de variation de deux variables aléatoires. Si les deux variables sont indépendantes, alors la covariance est nulle. (La réciproque est fausse).

Exercice 4. ♪

Montrer que $V(X + Y) = V(X) + V(Y) + 2Cov(X, Y)$.

Définition 4.10 Soient X et Y deux variables aléatoires d'écart type non nul.

On appelle **coefficient de corrélation linéaire** des deux variables X et Y , le nombre réel

$$\rho(X, Y) = \frac{Cov(X, Y)}{\sigma(X)\sigma(Y)}.$$

Exercice 5. ♪

Montrer que $|\rho(X, Y)| \leq 1$.

II Variables composées

Soit X une variable aléatoire et ϕ une fonction de $\mathbb{R}$ dans $\mathbb{R}$. Alors $\phi(X) = \phi \circ X$ est aussi une variable aléatoire que l'on note dans la suite $Y = \phi(X)$. Si $X(\Omega) = \{x_1, x_2, \dots, x_n\}$, alors

$$Y(\Omega) = \phi(X)(\Omega) = \{\phi(x_1), \dots, \phi(x_n)\} = \{y_1, \dots, y_m\},$$

cet ensemble étant de cardinal inférieur ou égal à n (cela dépend de l'injectivité de ϕ). On note $I_j = \{i \in \llbracket 1, n \rrbracket, \phi(x_i) = y_j\}$, ce qui veut dire que l'on découpe $\llbracket 1, n \rrbracket$ en sous-ensembles sur lequel ϕ est constante. En utilisant la réunion disjointe qui en découle, on trouve

$$p(Y = y_j) = \sum_{i \in I_j} p(X = x_i).$$

Proposition 4.11 (Théorème de transfert) Soit X une variable aléatoire et ϕ une fonction de $\mathbb{R}$ dans $\mathbb{R}$. On a

$$E(\phi(X)) = \sum_{i=1}^n \phi(x_i) p(X = x_i).$$

Exercice 6. ♪

Démontrer cette formule.

En particulier, on a

Corollaire 4.12 Avec les conditions précédentes, si ϕ est affine du type $\phi(x) = ax + b$, alors

$$E(aX + b) = aE(X) + b, \quad V(aX + b) = a^2V(X) \quad \text{et} \quad \sigma(aX + b) = |a|\sigma(X).$$

Exercice 7. ♪

Faire la démonstration.

Définition 4.13 Soit X une variable aléatoire discrète (i.e. $X(\Omega) \subset \mathbb{N}$). On définit la **fonction génératrice** G_X par :

$$G_X(t) = \sum_{k \in \mathbb{N}} t^k p(X = k)$$

Si la variable aléatoire est discrète finie (i.e. $\text{card}(X(\Omega)) = n$), la fonction génératrice est un polynôme. Sinon, il s'agit d'une série.

Exercice 8. ♪

Si X est une variable aléatoire et si $X(\Omega) = \llbracket 0, n \rrbracket$, on considère G_X la fonction génératrice associée à X .

1) Montrer que $\forall i \in \llbracket 0, n \rrbracket, p(X = i) = \frac{G_X^{(i)}(0)}{i!}$.

2) a) Montrer que $E(X) = G_X'(1)$.

b) Montrer que $V(X) = G_X''(1) + G_X'(1) - (G_X'(1))^2$.

3) a) Quel théorème permet de dire que $G_X(t) = E(t^X)$.

b) Soient $p \geq 2$ un entier et (X_i) une famille de p variables aléatoires mutuellement indépendantes à valeurs dans $\llbracket 0, n \rrbracket$. Montrer que $G_{\sum_{i=1}^p X_i} = \prod_{i=1}^p G_{X_i}$.

III Lois usuelles

III.1 Loi uniforme : $\mathcal{U}(n)$

C'est le cas d'un tirage équiprobable parmi n objets, X étant le numéro de l'objet tiré. En supposant que $X(\Omega) = \{1, \dots, n\}$ et $p(X = k) = \frac{1}{n}$ pour tout $k \in \{1, \dots, n\}$. On a alors :

$$\begin{aligned} E(X) &= \frac{n+1}{2} \\ V(X) &= \frac{n^2-1}{12} \end{aligned}$$

III.2 Loi de Bernoulli : $\mathcal{B}(p)$

Il s'agit du cas très simple où l'expérience aléatoire n'a que deux issues possibles appelées succès (avec une probabilité p) et échec (avec une probabilité $q = 1 - p$).

On a alors $X(\Omega) = \{0, 1\}$ et

$$\begin{aligned} E(X) &= p \\ V(X) &= pq \end{aligned}$$

III.3 Loi binomiale : $\mathcal{B}(n, p)$

Ici, on réalise n épreuves aléatoires **identiques et indépendantes** de Bernoulli. On a alors un arbre de probabilité de valence 2 (i.e. deux branches de plus à chaque ramification). En appelant X la variable aléatoire qui recense le nombre de succès, on a $X(\Omega) = \{0, 1, \dots, n\}$ et on a la

Proposition 4.14 Avec les mêmes hypothèses

$$p(X = k) = \binom{n}{k} p^k q^{n-k}$$

et

$$\begin{aligned} E(X) &= np \\ V(X) &= npq \end{aligned}$$

Exercice 9. ♪

Soit n , a et b trois entiers naturels non nuls. On dispose de n sacs contenant chacun a jetons numérotés $+1$ et b jetons numérotés -1 . On tire au hasard un jeton dans chaque sac. Soit T la somme des n valeurs relevées sur les jetons tirés.

- 1) Donner la loi de la variable aléatoire T .
- 2) Calculer son espérance mathématique et sa variance.

Exercice 10. ♪

La somme de deux variables aléatoires indépendantes X_1 et X_2 suivant deux lois binomiales $\mathcal{B}(n_1, p)$ et $\mathcal{B}(n_2, p)$ suit une loi binomiale $\mathcal{B}(n_1 + n_2, p)$.

III.4 Loi Multinomiale.

Si l'on reprend l'exercice classique de Babar :

Exercice 11. ♪

Un éléphant pioche cinq lettres avec remise dans un sac constitué des lettres B, A et R. Quelle est la probabilité que l'on puisse écrire "BABAR" avec les lettres tirées ?

On peut partir sur un dénombrement classique d'un tirage avec remise et avec ordre. On a alors le cardinal de l'univers qui vaut : $\text{card } \Omega = 3^5$.

Le nombre de tirages favorables est le nombre de codes à cinq lettres écrits avec trois lettres, dont deux fois le B, deux fois le A et une seule fois le R. Il suffit donc de choisir les places, soit

$$\binom{5}{2} \times \binom{3}{2} \times \binom{1}{1} = \frac{5!}{2!3!} \times \frac{3!}{2!1!} \times \frac{1!}{1!0!} = \frac{5!}{2!2!1!} = 30$$

Reprenons le même problème mais avec le point de vue algébrique : Si l'on veut développer le trinôme du cinquième degré $(b + a + r)^5$, on va obtenir une somme de monômes tous de degré global en a, b, r égal à 5. En effet, l'écriture devant être homogène, on doit avoir

$$(b + a + r)^5 = \sum_{\alpha_1 + \alpha_2 + \alpha_3 = 5} \text{Coeff} b^{\alpha_1} a^{\alpha_2} r^{\alpha_3}$$

On peut essayer de déterminer les coefficients. Au hasard, pour $b^2 a^2 r^1$, on a le développement suivant :

$$(b + a + r)^5 = (b + a + r) \times (b + a + r) \times (b + a + r) \times (b + a + r) \times (b + a + r) = \dots + 30b^2a^2r^1 + \dots$$

à vous de colorier les branches manquantes... et on retrouve bien notre nombre 30 trouvé précédemment. En effet, on a bien cherché le nombre de mots que l'on peut écrire avec deux B, deux A et un R, autrement dit, le nombre de fois où on a écrit $b^2 a^2 r$ dans le développement.

On peut évidemment étendre cette formule pour obtenir la formule du coefficient multinomial : $\frac{n!}{\alpha_1! \alpha_2! \dots \alpha_k!}$ lorsque l'on développe $(a_1 + \dots + a_k)^n$. Ainsi, on obtient la superbe formule :

$$(a_1 + \dots + a_k)^n = \sum_{\alpha_1 + \dots + \alpha_k = n} \frac{n!}{\alpha_1! \alpha_2! \dots \alpha_k!} a_1^{\alpha_1} \dots a_k^{\alpha_k}$$

qui généralise la formule du binôme de Newton.

Ceci nous permet de définir une nouvelle loi (très peu utilisée car on la retrouve par dénombrement à chaque fois sans problème), la loi multinomiale qui généralise la loi binomiale dans le cas où l'expérience aléatoire a plus de deux issues possibles (trois dans notre exemple).

III.5 Loi hypergéométrique : $\mathcal{H}(N, n, p)$

Dans une urne contenant N boules dont N_1 blanches et N_2 noires, on tire au hasard $n \leq N$ boules, sans remise. On appelle $p = \frac{N_1}{N}$ la proportion de boules blanches dans l'urne.

Soit X la variable aléatoire du nombre de boules blanches. Ses valeurs sont comprises entre 0 et n . En effet, la valeur maximum de X est n , sauf si $N_1 < n$, auquel cas, c'est N_1 . D'autre part, la valeur minimum de X est 0, sauf s'il n'y a pas assez de boules noires, i.e. si $N_2 < n$, auquel cas le minimum est $n - N_2$. Ainsi, on a $X(\Omega) = [\max\{0, n - N_2\}, \min\{n, N_1\}]$.

L'univers est un ensemble à $\binom{N}{n}$ éléments et pour tout $k \in X(\Omega)$, on prend k boules blanches parmi N_1 et $n - k$ parmi les N_2 boules noires. Ainsi, la probabilité de l'événement $(X = k)$ est :

Proposition 4.15 Si X suit une loi hypergéométrique

$$p(X = k) = \frac{\binom{N_1}{k} \binom{N_2}{n-k}}{\binom{N}{n}}$$

Remarque: Avec la notation $\binom{p}{l} = 0$ si $p < l$, on peut écrire :

$$1 = \sum_{k=0}^n p(X = k) = \sum_{k=0}^n \frac{\binom{N_1}{k} \binom{N_2}{n-k}}{\binom{N}{n}}$$

donc

$$\sum_{k=0}^n \binom{N_1}{k} \binom{N_2}{n-k} = \binom{N}{n}$$

On retrouve l'**Identité de Vandermonde**.

On a alors :

Proposition 4.16 Avec les mêmes hypothèses,

$$\begin{aligned} E(X) &= np \\ V(X) &= npq \left(\frac{N-n}{N-1} \right) \end{aligned}$$

Remarque: Le nombre $\left(\frac{N-n}{N-1} \right)$ est appelé **coefficient d'exhaustivité**.

On remarque que si N est grand ($N \geq 50$) et $\frac{n}{N}$ petit ($\frac{n}{N} < 0,1$), alors $\frac{N-n}{N-1} \simeq 1$, donc $\mathcal{H}(N, n, p) \simeq \mathcal{B}(n, p)$.

Démonstration: On décide de choisir les arrangements car les calculs vont être plus faciles qu'avec les combinaisons.

Notons

$$X_i = \begin{cases} 1 & \text{si la } i^{\text{e}} \text{ case comporte un élément du type recherché (boule blanche pour notre exemple)} \\ 0 & \text{sinon.} \end{cases}$$

Ainsi, $X_i \sim \mathcal{B}\left(\frac{n_1}{N}\right)$ i.e. $X_i \sim \mathcal{B}(p)$.

On a alors $X = \sum_{i=1}^n X_i$ et

$$\begin{aligned} E(X) &= \sum_{i=1}^n E(X_i) = np = \frac{nn_1}{N} \\ V(X) &= \sum_{i=1}^n V(X_i) + 2 \sum_{1 \leq i < j \leq n} \text{Cov}(X_i, X_j). \end{aligned}$$

Or

$$\begin{aligned} E(X_i \times X_j) &= 1 \times P(X_i = X_j = 1) + 0 \times P(X_i = 0 \text{ ou } X_j = 0) \\ &= P(X_i = X_j = 1) \\ &= \frac{A_{n_1}^2 A_{N-2}^{n-2}}{A_N^n} \\ &= \frac{n_1(n_1-1)}{N(N-1)}. \end{aligned}$$

Donc $\text{Cov}(X_i, X_j) = \frac{n_1(n_1-1)}{N(N-1)} - \left(\frac{n_1}{N}\right)^2 = \frac{-n_1 n_2}{N^2(N-1)}$ et

$$\begin{aligned} V(X) &= npq + 2 \binom{n}{2} \left(\frac{n_1}{N} \right) \\ &= npq - n(n-1) \frac{pq}{N-1} \\ &= npq \left(\frac{N-n}{N-1} \right) \end{aligned}$$

**Exercice 12.** ♪

Je joue au poker menteur. Je distribue cinq cartes à mon adversaire qui annonce aussitôt : "Brelan de roi !" Quel est la probabilité qu'il mente ?

- 1) S'il n'y a pas de joker dans le jeu (32 cartes).
- 2) S'il y a un joker dans le jeu ($32 + 1$ cartes et possibilité de mettre le joker dans le brelan).

