

Variables aléatoires discrètes

Outils Mathématiques & Statistique Décisionnelle

Licence 3 Management, Economie de la firme, CCF

Faculté d'économie, gestion & AES

Université de Bordeaux - Collège DSPEG

Automne 2024

A. Lourme

alexandre.lourme@u-bordeaux.fr

Outline

- Généralités
- Lois discrètes usuelles

Définitions

Une *variable aléatoire* X associe aux événements d'une expérience aléatoire d'univers Ω , une quantité d'intérêt : gain, perte, comptage, taille, etc¹. L'ensemble des valeurs prises par X est le *support* de X : $X(\Omega)$. X est une *variable aléatoire discrète* (v.a.d.) si $X(\Omega)$ est fini ou dénombrable.

La *fonction de probabilité* f d'une v.a.d. X associe à chaque valeur $x \in X(\Omega)$ la probabilité $f(x)$ de l'événement : $X = x$; et à $x \notin X(\Omega)$ la probabilité $f(x) = 0$.

La *fonction de répartition* F d'une v.a.d. X associe à chaque valeur $x \in \mathbb{R}$ la probabilité $F(x)$ de l'événement : $X \leq x$.

X étant discrète sa fonction de probabilité est représentée par un diagramme en bâtons et sa fonction de répartition par une fonction en escalier.

La fonction de répartition et la fonction de probabilité d'une v.a.d. X sont liées par :

$$\forall x \in \mathbb{R}, f(x) = F(x) - F(x^-) \quad (1)$$

où $F(x^-)$ est la limite de F à gauche de x . (1) permet de déduire la fonction de probabilité de la fonction de répartition, et réciproquement.

¹ cette définition, présentant une variable aléatoire comme une simple application, est incomplète mais on s'en contentera pour cette année

Exemple

M. Li lance un dé équilibré à six faces et gagne : 2 € si la face apparente est multiple de 2 ; 3 € si la face apparente est multiple de 3 ; 0 € sinon.

X : le gain de M. Li. L'univers : $\Omega = \{1, 2, 3, 4, 5, 6\}$. Le support de X :

$$X(\Omega) = \{0, 2, 3, 5\}.$$

$$\mathbb{P}(X = 0) = \mathbb{P}(\text{face ap.} \in \{1, 5\}) = 1/3 ; \mathbb{P}(X = 3) = \mathbb{P}(\text{face ap.} = 3) = 1/6$$

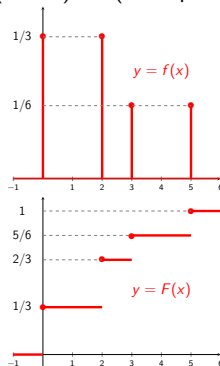
$$\mathbb{P}(X = 2) = \mathbb{P}(\text{face ap.} \in \{2, 4\}) = 1/3 ; \mathbb{P}(X = 5) = \mathbb{P}(\text{face ap.} = 6) = 1/6$$

Fonction de probabilité de X

$$f(x) = \begin{cases} 1/3 & \text{si } x \in \{0, 2\} \\ 1/6 & \text{si } x \in \{3, 5\} \\ 0 & \text{sinon} \end{cases}$$

Fonction de répartition de X

$$F(x) = \begin{cases} 0 & \text{si } x < 0 \\ 1/3 & \text{si } 0 \leq x < 2 \\ 2/3 & \text{si } 2 \leq x < 3 \\ 5/6 & \text{si } 3 \leq x < 5 \\ 1 & \text{si } x \geq 5 \end{cases}$$



Espérance & variance

X est une v.a.d. de fonction de probabilité f .

L'*espérance* de X : $E(X) = \sum_{x \in X(\Omega)} x.f(x)$ représente la valeur attendue de X .

La *variance* de X est :

$$V(X) \stackrel{(a)}{=} E((X - E(X))^2) \stackrel{(b)}{=} E(X^2) - E(X)^2 \stackrel{(c)}{=} \sum_{x \in X(\Omega)} x^2.f(x) - E(X)^2.$$

(a) $V(X)$ est l'écart quadratique moyen entre X et sa valeur attendue.

(b) Propriété de Koenig-Huygens

(c) $E(g(X)) = \sum_{x \in X(\Omega)} g(x).f(x)$ (Théorème de transfert)

Aspects informatiques.

Déterminer l'espérance et la variance d'une variable discrète passe par le calcul de sommes finies ou infinies (séries).

Sous R, la loi de probabilité du gain de M. Li (Slide 4) est donnée par :
`gain=c(0,2,3,5) ; proba=c(1/3,1/3,1/6,1/6)`

L'espérance du gain s'obtient par : `sum(proba*gain)`

Et la variance par : `sum(proba*gain^2)-sum(proba*gain)^2`

Outline

- Généralités
- Lois discrètes usuelles

Loi uniforme

Une v.a.d. X est distribuée selon une loi uniforme sur l'intervalle d'entiers $\llbracket 1; n \rrbracket$ si elle prend pour valeurs chacun des entiers $1, 2, \dots, n$ avec probabilité $1/n$.

Notation. $X \sim \mathcal{U}_{\llbracket 1; n \rrbracket}$

Espérance et variance. $E(X) = (1 + n)/2$; $V(X) = (n^2 - 1)/12$.

Contexte. Une urne contient n boules numérotées de 1 à n indiscernables au toucher. Si l'on choisit une boule de l'urne au hasard, le numéro tiré est distribué selon $\mathcal{U}_{\llbracket 1; n \rrbracket}$.

Exercice. Chaque jour, M. Li boit de un à quatre verres d'eau de façon équiprobable.

- (a) Tracez la fonction de probabilité et la fonction de répartition du nombre de verres d'eau absorbés par M. Li un jour donné.
- (b) Avec quelle probabilité M. Li boit-il au plus trois verres d'eau ?
- (c) Combien de verres d'eau M. Li boit-il par jour, en moyenne ?
- (d) Quelle est la variance du nombre quotidien de verres d'eau absorbés par M. Li ?

Loi de Bernoulli

Une v.a.d. X est distribuée selon une loi de Bernoulli de paramètre p ($0 < p < 1$) si elle prend pour valeurs : 1 avec probabilité p ou 0 avec probabilité $1 - p$.

Notation. $X \sim \mathcal{B}(p)$

Espérance et variance. $E(X) = p$; $V(X) = p \times (1 - p)$.

Contexte. Une expérience aléatoire comporte deux issues : le succès de probabilité p , ou l'échec. La variable X qui vaut 1 en cas de succès et 0 en cas d'échec est distribuée selon $\mathcal{B}(p)$.

Exercice 1. Quelle est la loi de X^2 lorsque $X \sim \mathcal{B}(p)$?

Exercice 2. Pour quelle valeur de p la variance de $X \sim \mathcal{B}(p)$ est-elle maximale ?

Loi binomiale

Une v.a.d. X est distribuée selon une loi binomiale de paramètres n et p ($n \in \mathbb{N} \setminus \{0\}; 0 < p < 1$) si elle prend toutes les valeurs entières entre 0 et n et pour $k \in \{0, 1, 2, \dots, n\}$, $\mathbb{P}(X = k) = \binom{n}{k} \times p^k \times (1 - p)^{n-k}$.

Notation. $X \sim b(n, p)$

Espérance et variance. $E(X) = n \times p$; $V(X) = n \times p \times (1 - p)$.

Contexte. Une expérience élémentaire comporte deux issues : le succès de probabilité p , ou l'échec. La variable X qui compte le nombre de succès dans la réalisation de n expériences élémentaires indépendantes est distribuée selon $b(n, p)$.

Exercice.

La classe de M. Li compte 20% de filles. M. Li envoie tour à tour cinq élèves au tableau. Avec quelle probabilité y-a-t-il (a) trois filles exactement (b) au plus trois filles (c) au moins trois filles parmi les cinq élèves interrogés ?

X : nombre de filles parmi les cinq élèves interrogés. $X \sim b(5; 0, 2)$.

$$(a) \mathbb{P}(X = 3) = \binom{5}{3} \times 0,2^3 \times 0,8^2 = 0,0512 \text{ (R : dbinom(3,5,0.2))}$$

$$(b) \mathbb{P}(X \leq 3) = \mathbb{P}(X \in \{0, 1, 2, 3\}) = \sum_{k=0}^3 \binom{5}{k} \times 0,2^k \times 0,8^{5-k} \approx 0,99 \text{ (R : pbinom(3,5,0.2))}$$

$$(c) \mathbb{P}(X \geq 3) = 1 - \mathbb{P}(X < 3) = 1 - \mathbb{P}(X \leq 2) = 0,05792 \text{ (R : 1-pbinom(2,5,0.2))}$$

Loi hypergéométrique

Une v.a.d. X est distribuée selon une loi hypergéométrique de paramètres N , n et p ($N \in \mathbb{N} \setminus \{0\}$; $n \in \{0, 1, \dots, N\}$; $0 < p < 1$) si :

- elle prend toutes les valeurs entières entre $\max\{0, n - N(1 - p)\}$ et $\min\{n, Np\}$
- k étant l'une de ces valeurs, $\mathbb{P}(X = k) = \binom{Np}{k} \times \binom{N(1-p)}{n-k} / \binom{N}{n}$.

Notation. $X \sim \mathcal{H}(N, n, p)$

Espérance et variance. $E(X) = n \times p$; $V(X) = n \times p \times (1 - p) \times (N - n) / (N - 1)$.

Contexte. Dans une population finie de taille N , une proportion p de sujets possède une qualité et l'on choisit n sujets simultanément. La variable X comptant le nombre de sujets de l'échantillon possédant la qualité étudiée est distribuée selon $\mathcal{H}(N, n, p)$.

Exercice.

La classe de M. Li compte 20% de filles sur 30 élèves. M. Li envoie cinq élèves au tableau simultanément. Avec quelle probabilité y-a-t-il (a) trois filles exactement (b) au plus trois filles parmi les cinq élèves interrogés ?

X : nombre de filles parmi les cinq élèves interrogés.

$X \sim \mathcal{H}(30; 5; 0, 2)$ (30 élèves dans la classe dont 20% de filles et 5 élèves interrogés).

$$(a) \mathbb{P}(X = 3) = \binom{6}{3} \binom{24}{2} / \binom{30}{5} \approx 0,039 \text{ (R : dhyper(3, 6, 24, 5))}$$

$$(b) \mathbb{P}(X \leq 3) = \mathbb{P}(X \in \{0, 1, 2, 3\}) = \sum_{k=0}^3 \binom{6}{k} \binom{24}{5-k} / \binom{30}{5} = 0,997 \text{ (R : phyper(3, 6, 24, 5))}$$

Loi géométrique

Une v.a.d. X est distribuée selon une loi géométrique de paramètres p ($0 < p < 1$) si elle prend pour valeurs les entiers non nuls et pour $k \in \mathbb{N} \setminus \{0\}$,
 $\mathbb{P}(X = k) = (1 - p)^{k-1}p$.

Notation. $X \sim \mathcal{G}(p)$

Espérance et variance. $E(X) = 1/p$; $V(X) = (1 - p)/p^2$.

Contexte. Une expérience aléatoire élémentaire comporte deux issues : le succès de probabilité p , ou l'échec. La variable X qui compte le nombre de tentatives jusqu'au premier succès est distribuée selon $\mathcal{G}(p)$.

Exercice.

Au basket, Bob a 60% de chances de réussir un panier. Avec quelle probabilité faudra-t-il à Bob (a) trois essais (b) au plus trois essais, jusqu'au premier panier ?

X : nbre d'essais jusqu'au premier panier. $X \sim \mathcal{G}(0,6)$.

(a) $\mathbb{P}(X = 3) = 0,4^2 \times 0,6 = 0,096$ (R : `dgeom(3-1,0.6)`)

(b) $\mathbb{P}(X \leq 3) = \mathbb{P}(X \in \{1, 2, 3\}) = \sum_{k=1}^3 0,4^{k-1} \times 0,6 = 0,94$ (R : `pgeom(3-1,0.6)`)

NB. R compte le nombre d'essais précédant le premier succès.

Loi de Poisson

Une v.a.d. X est distribuée selon une loi de Poisson de paramètres λ ($\lambda > 0$) si elle prend pour valeurs les entiers naturels et pour $k \in \mathbb{N}$, $\mathbb{P}(X = k) = e^{-\lambda} \lambda^k / k!$.

Notation. $X \sim \mathcal{P}(\lambda)$

Espérance et variance. $E(X) = \lambda$; $V(X) = \lambda$.

Contexte. Une variable qui compte le nombre d'occurrences d'un événement rare dans un intervalle de temps/de longueur/de surface/... restreint est distribuée selon une loi de Poisson.

Exercice.

Au football, le nombre de buts dans un intervalle de temps de vingt minutes est distribué selon $\mathcal{P}(1/2)$. Avec quelle probabilité observe-t-on (a) un but exactement (b) au plus un but, en vingt minutes ?

X : nbre de buts observés en vingt minutes. $X \sim \mathcal{P}(1/2)$.

(a) $\mathbb{P}(X = 1) = e^{-1/2} (1/2)^1 / 1! \approx 0,30$ (R : dpois(1,0.5))

(b) $\mathbb{P}(X \leq 1) = \mathbb{P}(X \in \{0, 1\}) = \sum_{k=0}^1 e^{-1/2} (1/2)^k / k! \approx 0,91$ (R : ppois(1,0.5))

Approximations

Approximation de la loi hypergéométrique par une loi binomiale

La loi $\mathcal{H}(N, n, p)$ peut être approchée par la loi $b(n, p)$ quand $N \geq 10n$.

Exemple.

Une promotion compte 500 étudiants dont 30% de filles. Avec quelle probabilité α un échantillon de quarante étudiants compte-t-il dix-huit filles ?

Sans approx. $\alpha = \mathbb{P}(\mathcal{H}(500; 40; 0, 3) = 18) = \binom{500 \times 0,3}{18} \binom{500 \times 0,7}{40-18} / \binom{500}{40} \approx 0,015$

Avec approx. $\alpha \approx \mathbb{P}(b(40; 0, 3) = 18) = \binom{40}{18} \times 0,3^{18} \times 0,7^{22} \approx 0,017$

Approximation de la loi binomiale par une loi de Poisson

La loi $b(n, p)$ peut être approchée par la loi $\mathcal{P}(\lambda = n \times p)$ quand $n \geq 30$ et $p \leq 0,1$.

Exemple.

10% des français sont favorables à M. Sa. Avec quelle probabilité α un échantillon de quarante français compte-t-il cinq partisans de M. Sa ?

Sans approx. $\alpha = \mathbb{P}(b(40; 0, 1) = 5) = \binom{40}{5} \times 0,1^5 \times 0,9^{35} \approx 0,165$

Avec approx. $\alpha \approx \mathbb{P}(\mathcal{P}(40 \times 0, 1) = 5) = e^{-4} \times 4^5 / 5! \approx 0,156$