Variables aléatoires discrètes

Rappels

Variables aléatoires

Une *variable aléatoire* (V.A.) est une fonction à valeurs numériques, définie sur un espace Ω. Soit $\mathcal{F}$ un phénomène fortuit et Ω son espace (ensemble des résultats possibles de $\mathcal{F}$), alors la variable aléatoire X est définie comme la fonction

$$\forall \omega \in \Omega : \omega \mapsto X(w) = x \in \mathbb{R}$$

La terminologie adaptée est trompeuse : une variable aléatoire n'est pas une variable au sens où on l'entend habituellement en analyse mais une application. De plus, elle n'est pas aléatoire. C'est l'expérience qui est aléatoire.
On appelle *fonction de répartition* de X la fonction F définie sur $\mathbb{R}$ par

$$\forall x \in \mathbb{R} : F(x) = \mathbb{P}(X \leqslant x)$$

Une variable aléatoire X est *discrète* si l'ensemble de valeurs prises par X est dénombrable.
Dans le cas discrèt, on appelle *espérance mathématique* de X le nombre $E[X]$ défini par

$$E[X] = \sum_{n=0}^{\infty} x_n \mathbb{P}(X = x_n)$$

On appelle *variance* de X le nombre $V[X]$ défini par

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

<u>Propriétés de l'espérance et de la variance</u>

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

$$E[aX + b] = aE[X] + b$$
$$V[aX + b] = a^2V[X]$$

Loi discrète uniforme

Une variable aléatoire X suit une *loi uniforme discrète* sur l'intervalle $[1, n]$ de $\mathbb{N}$ si

$$X(\Omega) = \{1, \ldots, n\} \quad \text{et} \quad \mathbb{P}(X = k) = \frac{1}{n}.$$

On choisit au hasard (i.e. avec équiprobabilité) un objet parmi n.

$$E[X] = \frac{n+1}{2} \quad \text{et} \quad V[X] = \frac{n^2 - 1}{12}$$

Loi de Bernoulli

Une variable aléatoire X suit une *loi de Bernoulli* si elle ne prend que les deux valeurs 0 et 1 avec des probabilités non nulles. On note souvent $\mathbb{P}(X = 1) = p$ et $\mathbb{P}(X = 0) = 1 - p = q$, le résultat 1 étant considéré comme un succès et 0 comme un échec.

Une variable de Bernoulli illustre donc toute expérience aléatoire n'ayant que deux issues possibles et effectuée une seule fois.

$$E[X] = p \quad \text{et} \quad V[X] = pq$$

Loi Binomiale

Une variable aléatoire X est distribuée selon une *loi Binomiale* de signature (n, p) avec $n \in \mathbb{N}$ et $p \in [0, 1]$ si elle correspond au nombre de succès obtenus lors d'une succession de n essais indépendants et identiques d'une expérience aléatoire n'ayant que deux résultats possibles : le succès avec une probabilité p et l'échec avec une probabilité $q = 1 - p$. On note $X \sim B(n, p)$.
Si $X \sim B(n, p)$ alors $\forall k \in \{1, \ldots, n\}$

$$\mathbb{P}(X = k) = \mathcal{C}_n^k p^k q^{n-k}.$$

Une variable aléatoire binomiale de signature (n, p) peut être vue comme la loi de probabilité caractérisant n tirages avec remise hors d'une urne contenant des boules d'une certaine couleur en proportion p et d'une autre couleur en proportion $1 - p$.

$$E[X] = np \quad \text{et} \quad V[X] = npq$$

Loi de Poisson

Une variable aléatoire X suit une *loi de Poisson* de paramètre $\lambda > 0$, $X \sim P(\lambda)$, si $X(\Omega) = \mathbb{N}$ et

$$\mathbb{P}(X = k) = e^{-\lambda}\frac{\lambda^k}{k!} \forall k \in \mathbb{N}.$$

Cette loi est souvent utilisé dans la pratique pour dénombrer les nombres d'occurences d'un phénomène bien précis dans un intervalle de temps doné. Citons, par exemple, le nombre d'appels téléphoniques aboutissant à un standard durant 5 minutes, le nombre de clients se présentant à un guichet de poste entre 14h et 16h, le nombre d'accidents observés sur une autoroute pendant une heure de pointe,...

$$E[X] = \lambda \quad \text{et} \quad V[X] = \lambda$$

Approximation de la loi Binomiale par la loi de Poisson :
Une loi de Poisson de paramètre $\lambda = np$ est une bonne approximation de la loi Binomiale de signature (n, p) quand n est grand et p petit ($n \geqslant 30$, $p \leqslant 0,1$ et $np \leqslant 10$).

Loi géométrique

Une variable aléatoire X suit une *loi géométrique* si $X(\Omega) = \mathbb{N}_0$ et

$$\mathbb{P}(X = k) = (1 - p)^{k-1}p$$

Reprenons le modèle probabiliste d'une variable aléatoire Binomiale. Au lieu de s'intéresser au nombre de succès, on peut vouloir déterminer le nombre de tirages nécessaires jusqu'à ce qu'une telle boule soit tirée.

$$E[X] = \frac{1}{p} \quad \text{et} \quad V[X] = \frac{1-p}{p^2}$$

Loi Binomiale négative

Une variable aléatoire X suit une *loi Binomiale négative* de signature (r, p) si elle correspond au nombre d'essais nécessaires afin d'obtenir r succès lors d'une succession d'essais indépendants et identiques d'une expérience aléatoire n'ayant que deux résultats possibles : le succès avec une probabilité p et l'échec avec une probabilité $q = 1 - p$. On a que $X(\Omega) = \{r, r+1, \ldots\}$ et

$$\mathbb{P}(X = k) = \mathcal{C}_{k-1}^{r-1}p^r q^{k-r}$$

Une variable aléatoire Binomiale négative de signature (r, p) peut être vue comme la loi de probabilité caractérisant le nombre de tirages avec remise hors d'une urne contenant des boules d'une certaine couleur en proportion p et d'une autre couleur en proportion $1-p$ nécessaires afin d'en avoir r de la première couleur.

$$E[X] = \frac{r}{p} \quad \text{et} \quad V[X] = \frac{r(1-p)}{p^2}$$

Loi Hypergéométrique

Une variable aléatoire X est distribuée selon une *loi Hypergéométrique* de signature (N, n, m) si elle correspond au nombre de boules blanches contenus dans un échantillon de taille n prélevé aléatoirement dans une urne contenant N boules dont m sont blanches et $N-m$ sont noires. On a que $X(\Omega) = \{0, 1, \ldots, n\}$ et

$$\mathbb{P}(X = k) = \frac{\mathcal{C}_m^k \mathcal{C}_{N-m}^{n-k}}{\mathcal{C}_N^n}$$

$$E[X] = \frac{nm}{N} \quad \text{et} \quad V[X] = np(1-p)(1 - \frac{n-1}{N-1})$$

Exercices

1. Un canal de transmission d'information ne peut traiter que des 0 et des 1. À cause de perturbations dues à l'électricité statique, chaque chiffre transmis l'est avec une probabilité d'erreur de $0,2$. Admettons que l'on veuille transmettre un message important limité à un signal binaire. Pour éviter une erreur, on transmettrea 00000 au lieu de 0 et 11111 au lieu de 1. Si le récepteur décode suivant la règle de la majorité, quelle est la probabilité que le message soit mal interprété ?
2. Un épicier reçoit un lot de pommes dont 25% sont avariés. Il charge un employé de préparer des emballages de 5 pommes chacun. Celui-ci, négligent, ne se donne pas la peine de jeter les fruits avariés. Chaque client qui trouve, dans l'emballage qu'il achète, 2 fruits ou plus qui sont avariés, revient au magasin se plaindre.
   a) Soit X le "nombre de pommes avariées dans un emballage". Déterminer la loi de probabilité de X.
   b) Quelle est la probabilité pour qu'un client donné se plaigne auprès de son épicier ?

c) Si l'épicier a 100 clients qui achètent des pommes ce jour-là, combien y aura-t-il de plaintes ?

3. Giovanni, dit Gianni, a décidé de parcourir 10000 km en Fiat Ritmo. Or la probabilité d'avoir un accident sur 1 km est de $1/10000$. En prenant connaissance de cette probabilité, Gianni décide alors d'annuler so long voyage en prétextant être sûr d'avoir un accident. Êtes-vous d'accord avec Gianni ? Quelle est alors approximativement la probabilité que Gianni ait un accident ?
4. Un journaliste se voit remettre une liste de personnes à interviewer. Il doit interroger 5 personnes au moins. Les interviewés potentiels n'acceptent de parler qu'avec une probabilité de $2/3$, indépendamment les uns des autres. Quelle est la probabilité qu'il puisse réaliser ses 5 entretiens si la liste compte 5 noms ? Et si elle en compte 8 ?
5. Des études effectuées par les compagnies aériennes montrent qu'il y a une probabilité de $0,05$ que chaque passager ayant fait une réservation n'effectue pas le vol. De ce fait, SA Airlines vend toujours 94 billets pour ses avions à 90 sièges, tandis que BA Airlines vend toujours 188 billets pour ses avions à 180 sièges. Avec quelle compagnie un passager ayant réservé un siège risque-t-il le plus de ne pas pouvoir prendre place dans l'avion ?
6. Madame Gourmande prépare des biscuits aux raisins secs et aux pépites de chocolat. Elle mélange 600 raisins et 400 pépites de chocolat dans la pâte et prépare 500 biscuits. Dès que les biscuits sont prêts, Madame Gourmande en choisit un au hasard pour le goûter.
   a) Calculer la probabilité qu'il n'y ait pas de raisins dans le biscuit.
   b) Calculer la probabilité qu'il y ait exactement 2 pépites de chocolat.
   c) Calculer la probabilité qu'il y ait au moins 2 morceaux (raisins ou pépites de chocolat).
7. a) Nadine part à la cueillette des champignons. Elle ne sait pas faire la différence entre un champignon comestible et un champignon toxique. On estime que la proportion de champignons toxiques se trouvant dans les bois s'élève à $0,7$.
      i. Nadine ramasse 6 champignons au hasard. Calculer la probabilité qu'elle ramasse exactement 4 champignons toxiques.
      ii. Nadine invite Serge à une ceuillette. Serge connaît bien les champignons ; sur 10 champignons qu'il ramasse, 9 sont comestibles. Ce jour-là, il ramasse 4 champignons et Nadine en ramasse 3. Calculer la probabilité que tous les champignons soient comestibles.

b) Serge cueille en moyenne 12 champignons par heure.

i. Calculer la probabilité qu'il ramasse exactement 8 champignons en une heure.

ii. Calculer la probabilité qu'il ramasse au moins 1 champignon en 20 minutes.

8. Les ingénieurs du son préposés à la sonorisation d'un concert en plein air hésitent entre deux solutions : 4 haut-parleurs de 4000 watts chacun ou 2 haut-parleurs de 8000 watts chacun. On suppose que la probabilité qu'un haut-parleur tombe en panne est égale à $p = 0,2$ indépendamment du type de haut-parleur et que les pannes se produisent indépendamment les unes des autres. En admettant que le concert peut se dérouler avec au moins 8000 watts, quelle solution conseillerez-vous à ces ingénieurs ?

9. On vous demande de déterminer l'efficacité d'un test de détection e la grippe aviaire sur les poulets. On suppose que si un poulet n'est pas infecté, le teste est positif avec une probabilité $0,02$. Lorsque ce dernier est infecté, le text est négatif avec probabilité $0,01$. On sait que 10% des poulets sont infectés.

   a) Le test est positif sur un poulet. Quelle est la probabilité que ce poulet soit infecté ?

   b) Sur un lot de dix poulets, deux sont testés positivement et le reste négativement. Quelle est la probabilité qu'aucun des dix poulets ne soit infecté ?

10. Chacun des soldats d'une troupe de 500 hommes est porteur d'une certaine maladie avec probabilité $1/1000$. Cette maladie est détectable à l'aide d'un test sanguin et, pour faciliter les choses, on ne teste qu'un mélange du sang des 500 soldats.

    a) Quelle est la probabilité que le test soit positif, indiquant par là qu'au moins une des personnes est malade ?

    b) On suppose que le test a été positif. Quelle est la probabilité que dans ce cas, plus d'une personne soit malade ?

11. À l'univesité, le taux de blocage de mâchoire provoqué par un bâillement profond est de 1 personne pour 1000 et par mois. On suppose qu'un étudiant a au plus un blocage de mâchoire par mois et que les blocages de mâchoire sont indépendants.

    a) Quelle est la probabilité qu'en un mois de l'année académique $2005 - 2006$ il y ait 3 blocages de mâchoire ou plus parmi les étudiants de l'université ?
    Indication : supposer qu'il y a 4000 étudiants.

b) Quelle est la probabilité qu'au cours de l'année académique 2005 − 2006 le nombre de mois comptant 3 blocages de mâchoire ou plus soit supérieur ou égal à 4 ?

Indication : supposer qu'une année académique est constituée de 8 mois.

c) Le premier mois de l'année académique étant appelé mois 1, quelle est la probabilité que le premier mois où l'on registre 3 blocages de mâchoire ou plus soit le mois i, $i = 1, 2, \ldots, 8$?

12. Une entreprise aimerait lancer un nouveau produit sur le marché. Pour ce faire, l'entreprise va utiliser ses statistiques de marketing d'un produit similaire. Elle trouve que 3/4 des personnes dans la population cible ayant été exposés à une campagne publicitaire pour l'ancien produit affirment l'avoir acheté, tandis que 4/5 des personnes qui n'ont pas vu de publicité disent ne pas l'avoir acheté. 2/3 des personnes interrogées affirment avoir vu la publicité. On choisit aléatoirement un échantillon de 30 personnes et l'entreprise décide que le nouveau produit peut être lancé sur le marché si plus de 95% de ce groupe veut l'acheter.

    a) Calculer la probabilité que l'individu veut acheter le produit.

    b) Un individu tiré au hasard a affirmé vouloir acheter le produit. Quelle est la probabilité qu'il ait vu la publicité ? Généraliser le résultat pour k individus.

    c) Donner la distribution de nombre de personnes dans l'échantillon qui veulent acheter le produit et calculer la probabilité que le produit soit lacé sur le marché.

13. Une compagnie d'assurance propose à ces clients un nouveau contrat annuel destiné à les couvrir en cas de vol dans leur résidence secondaire. Sur base d'une étude statistique, la compagnie estime que six résidences secondaires sur 100 sont cambriolées au moins une fois par an et que la valeur des biens volés dépasse 2500 euros une fois sur 50. Supposons que 10 clients prennent l'assurance proposée.

    a) Quelle est la probabilité que la compagnie ne doive pas intervenir ?

    b) Quelle est la probabilité qu'elle doive intervenir pour deux clients ?

    Supposons que 100 clients prennent l'assurance proposée.

    c) Donnez une prévision du nombre de clients qui bénificeront de l'intervention de l'assurance.

    d) Quelle est la probabilité que l'assurance doive intervenir pour 3 clients au moins ? (Faites une approximation)

14. On suppose que le nombre d'appels parvenant à la police le samedi matin entre 6h et 6h30 suit une loi de Poisson dont le paramètre λ vaut $2,5$.
    a) Quelle est la probabilité qu'aucun appel ne parvienne à la police durant cette période ?
    b) Quelle est la probabilité d'avoit 3 appels ou plus ?
    c) Calculer l'espérance et la variance du nombre d'appels.
15. **Le problème des allumettes de Banach**
    Un mathématicien se trouve être également fumeur de pipe et il porte à tout moment deux boîtes d'allumettes, une dans chacune de ses poches gauche et droite. Chaque fois qu'il a besoin d'une allumette, il a une chance sur deux d'aller la chercher dans sa poche gauche et pareil pour la poche droite. Il découvre subitement que la boîte tirée est vide. Les deux boîtes contenaient au départ N allumettes chacune. Quelle est la probabilité qu'il lui reste k allumettes dans l'autre boîte, $k = 0, 1, \ldots, N$?
16. On considère l'expérience qui consiste à mesurer le nombre de particules α émises dans l'espace d'une seconde par un gramme de matière radioactive. Des expériences ont montré dans le passé qu'en moyenne le nombre de particules α émises est $3,2$. Donnez une bonne approximation pour la probabilité qu'au plus deux particules α seront enregistrées.
17. On suppose que les réacteurs d'un avion ont, avec une probabilité $1-p$, une défaillance en cours de vol. Les défaillances se produisent indépendamment les unes des autres. L'avion peut terminer son vol sans difficulté si la moitié de ses réacteurs au moins fonctionnent. Pour quelles valeurs de p les quadriréacteurs sont-ils préférables aux biréacteurs ?
18. Une population compte 1% de personnes diabétiques. On prélève, au hasard et sans remise, un échantillon de 20 personnes. Soit X la variable "nombre de diabétiques dans l'échantillon".
    a) Quelle est la loi de X ?
    b) Calculez $E[X]$ et $V[X]$.
    c) Calculez $\mathbb{P}(X > 0)$ et $\mathbb{P}(X > 2)$.
19. Le fil d'un métier à tisser se casse en moyenne $1,375$ fois par heure de fonctionnement du métier. Calculez la probabilité pour que durant huit heures de travail, le nombre X de cassures du fil se trouve entre 2 et 4.

20. La probabilité qu'une personne quelconque soit allergique est de $0,1\%$.
    a) Quelle est la loi de probabilité du nombre d'allergiques sur une population de 4000 personnes ?
    b) Peut-on réaliser une approximation de cette loi par une loi de Poisson ?
    c) Calculez la probabilité que le nombre d'allergiques ne dépasse pas deux personnes.
    d) Calculez la probabilité qu'il soit au moins de trois.
21. Un texte est donné à dactylographier en deux parties de 1000 caractères chacune à deux secrétaires différentes. La probabilité que la première sécrétaire fasse une erreur de frappe est de $0,002$ par caractère, elle est de $0,001$ pour la seconde secrétaire. Quelle est la probabilité de trouver dans le texte 3 erreurs au moins ? (Faites une approximation)
22. Une sonde posée sur Mars observe les impacts des météorites qui se produisent dans un rayon de 10 kilomètres autour d'elle. Le nombre d'impacts par jour martien suit une loi de Poisson de paramètre λ.
    a) Pendant 50 jours martiens consécutifs, la sonde a détecté en tout 150 impacts. Quelle valeur plausible pour λ peut-on choisir ? (On considère cette valeur particulière de λ dans la suite de l'exercice).
    b) Quelle est la probabilité qu'il n'y ait aucun impact durant une journée martienne ?
    c) Supposons à présent que sur 100 météorites qui tombent dans la zone d'observation, il y en ait un qui pèse plus d'une tonne.
        i. Sachant qu'il tombe k météorites un certain jour martien dans la zone d'observation, quelle est la loi du nombre de météorites qui pèsent plus d'une tonne ?
        ii. Quelle est la probablitié d'observer deux météorites de plus d'une tonne, sachant qu'il est tombé trois météorites durant la journée considérée ?
        iii. Quelle est la probabilité d'observer l'arrivée de trois météorites dont deux de plus d'une tonne dans la journée considérée ?
23. Une urne contient N boules blanches et M boules noires. Si les boules sont tirées de manière aléatoire avec remise l'une après l'autre jusqu'à ce qu'une boule noire soit tirée. Quelle est la probabilité que
    a) l'on nécessite exactement n tirages ?
    b) l'on nécessite au moins k tirages ?

24. Un marchand de pièces électriques achète ses pièces par des lots de 10. Il choisit aléatoirement 3 pièces d'un lot, les inspecte et achète le lot si elles sont toutes les trois fonctionnelles. Si 30% des lots ont 4 pièces défectueuses et 70% des lots ont une pièce défectueuse quelle est la proportion des lots rejetés ?
25. Un jeu de dés très populaire dans les bars anglais est le *chuck-a-luck*. Il consiste pour la banque à jeter 3 dés. Un joueur peut parier sur n'importe quel résultat compris entre 1 et 6. Si exactement un de ces 3 dés montre le chiffre prévu, le joueur récupère sa mise plus un montant équivalent. Si 2 dés montrent ce résultat, le gain net est de 2 pour 1 ; si les 3 dés indiquent le chiffre prévu, le gain net est de 3 pour 1. Si aucun dé ne montre le chiffre choisi par le joueur, ce dernier perd sa mise.
    a) Calculer l'espérance de gain lorsque l'enjeu est d'une livre.
    b) Quel montant devrait recevoir le joueur si les 3 dés montrent le chiffre prévu pour que le jeu soit *fair* (c'est-à-dire pour que l'espérance de gain soit nulle) ?
26. a) Soit une urne contenant 30 boules dont 15 sont blanches et 15 sont noires. Quelle est la probabilité, lors d'une succession de tirages indépendants avec remise, que l'on a tiré 10 boules blanches avant d'en avoir tiré 5 noires.
    b) Soit maintenant une urne contenant une proportion p de boules blanches. Quelle est la probabilité que l'on tire r boules blanches avant d'en avoir tiré m noires.
27. En moyenne, combien de fois faut il lancer un dé pour avoir obtenu 4 fois le chiffre 6 ? Ce résultat est-il fiable ?
28. Une urne contient N boules blanches et M boules noires. Quelle est la probabilité que n boules blanches sont tirées avant m boules noires si les boules soient tirées de façon indépendante, l'une après l'autre et avec remise ? ($n \leqslant N$ et $m \leqslant M$)
29. Une urne contient 4 boules blanches et 4 boules noires. On choisit 4 boules au hasard. S'il y a 2 boules blanches et 2 boules noires, on s'arrête. Sinon, on remet les boules dans l'urne et on recommence. Quelle est la probabilité que l'on recommence n fois ?
30. Supposons qu'un lot de 100 ampoules en contient 6 qui sont défectueuses et 94 qui ne sont pas défectueuses. Quelle est la probabilité que dans un échantillon de 10 ampoules
    a) aucune ne soit défectueuse ?
    b) plus de deux soient défectueuses ?