1

CHAPITRE 4

VARIABLES ALÉATOIRES DISCRÈTES

On considère tout au long de ce chapitre un univers, souvent noté Ω, qui est constitué de toutes les « éventualités » ou « issues » d'une expérience aléatoire.

4.1 Généralités sur les variables aléatoires discrètes

Définition 4.1 (Variable aléatoire discrète) *On appelle variable aléatoire discrète toute fonction :*

$$\begin{array}{rcl} X:\Omega & \to & E \\ \omega & \mapsto & X(\omega) \end{array}$$

où E est un ensemble fini ou dénombrable.

Notation 4.2 *Tout événement de la forme : $\{\omega \in \Omega \backslash X(\omega) \in A\}$ est noté $(X \in A)$ ou $[X \in A]$ pour tout $A \subset \mathbb{R}$. Ce qui se lit « l'ensemble des ω tels que $X(\omega)$ appartient à A ».*

Définition 4.3 (Loi de probabilité) *Soit X une variable aléatoire discrète.(abrégée par v.a.d.)*
La loi de probabilité de X est déterminée par :

- *l'ensemble des valeurs prises par X c-à-d $X(\Omega) = E = \{e_i, i \in I\}$ où I est fini ou dénombrable,*
- *les probabilités associées aux valeurs prises par X : $\{\mathbb{P}(X = e_i); i \in I\}$.*

Remarque 4.1 *En notant p_i les probabilités $\mathbb{P}(X = e_i)$, les p_i doivent satisfaire les deux conditions suivantes :*

1. $\forall i \in I, 0 \leqslant p_i \leqslant 1$,
2. $\displaystyle\sum_{i \in I} p_i = 1$.

Exemple 4.1 *On lance deux fois de suite un dé. Soit X la variable aléatoire discrète représentant le nombre de 6 obtenus. Déterminons la loi de X.*
On a : $\Omega = [\![1, 6,]\!] \times [\![1, 6]\!] = [\![1, 6,]\!]^2$ de cardinal $6^2 = 36$.
Ici X est à valeurs dans $X(\Omega) = E = \{0, 1, 2\}$.
Donner la loi de X consiste à déterminer : $\mathbb{P}(X = 0)$, $\mathbb{P}(X = 1)$ et $\mathbb{P}(X = 2)$.
On remarque d'après la définition que pour tout $k \in E$ on a :

$$\mathbb{P}(X = k) = \mathbb{P}(\{(u, v) \in \Omega | X(u, v) = k\}).$$

⊗ un ensemble est dits denombrable qu'on peut le metre en bijection avec IN .

⊗ Pour connaitre une loi de probabilité discrète, il suffit de connaitre les valeurs prises par la variable discrète et les probabilites correspendents .

ⓔ

Calculons ces probabilités :

$$\begin{aligned}
(X=0) &= [\![1,5]\!] \times [\![1,5]\!] \\
\Rightarrow \mathbb{P}(X=0) &= \frac{Card(X=0)}{Card(\Omega)} = \frac{25}{36} \\
(X=1) &= \{(1,6),(2,6)\ldots(5,6),(6,1),(6,2),\ldots(6,5)\} \\
\Rightarrow \mathbb{P}(X=1) &= \frac{Card(X=1)}{Card(\Omega)} = \frac{10}{36} \\
(X=2) &= \{(6,6)\} \\
\Rightarrow \mathbb{P}(X=2) &= \frac{Card(X=2)}{Card(\Omega)} = \frac{1}{36}
\end{aligned}$$

On remarque que : $\mathbb{P}(X=0)+\mathbb{P}(X=1)+\mathbb{P}(X=2)=1$.
C'est normal car les événements sont incompatibles entres eux :
$\mathbb{P}(X=0)+\mathbb{P}(X=1)+\mathbb{P}(X=2)=\mathbb{P}((X=0)\cup(X=1)\cup(X=2))=\mathbb{P}(\Omega)=1.$

4.1.0.1 Intérêt de connaître la loi d'une variable aléatoire

Si on connaît la loi de X, alors il est facile de déterminer la probabilité de tout événement ne dépendant que de X.

Exemple 4.2 *On lance deux dés discernables, et notons S la somme obtenue, après jet des deux dés. On a :* $\Omega = [\![1,6,]\!]^2$ *de cardinal* 36.
Ici S est à valeurs dans $E=[\![2,12]\!]$.
Intéressons-nous à déterminer par exemple la probabilité de l'événement $(S \leqslant 4)$.

$$\mathbb{P}(S \leqslant 4) = \mathbb{P}((S=2)\cup(S=3)\cup(S=4))$$

Or les événements sont incompatibles, donc :

$$\begin{array}{rrcl}
& \mathbb{P}(S \leqslant 4) & = & \mathbb{P}(S=2)+\mathbb{P}(S=3)+\mathbb{P}(S=4) \\
Or & (S=2) & : & \{(1,1)\} \\
& (S=3) & : & \{(1,2),(2,1)\} \\
& (S=4) & : & \{(1,3),(3,1),(2,2)\} \\
Donc & \mathbb{P}(S \leqslant 4) & = & \frac{1}{36}+\frac{2}{36}+\frac{3}{36}=\frac{6}{36}=\frac{1}{6}
\end{array}$$

4.2 Lois discrètes usuelles

4.2.1 Variable de Bernoulli

Définition 4.4 *On dit qu' une variable aléatoire X suit une loi* de Bernoulli de paramètres p *si :*

1. $X(\Omega)=\{0,1\}$,
2. $\mathbb{P}(X=1)=p, \mathbb{P}(X=0)=1-p=q$.

Notation 4.5 *On notera ceci en écrivant :* $X \hookrightarrow \mathcal{B}(p)$.

Exemple 4.3 *Une urne contient deux boules rouges et trois boules vertes. On tire une boule de l'urne. La variable aléatoire X = nombre de boules rouge(s) tirée(s) est une variable de Bernoulli. On a :* $\mathbb{P}(X=1)=2/5=p$, $\mathbb{P}(X=0)=3/5=q$.

Plus généralement, on utilisera une variable de Bernoulli lorsqu'on effectue une épreuve qui n'a que deux issues : le succès ou l'échec. Une telle expérience est alors appelée épreuve de Bernoulli. On affecte alors 1 à la variable en cas de succès et 0 en cas d'échec.

Remarque 4.2

$$\sum_{k=0}^{1} \mathbb{P}(X=k) = p+q = 1$$

4.2.2 Distribution binomiale

4.2.2.1 Situation concrète

a) On effectue une épreuve de Bernoulli. Elle n'a donc que deux issues : le succès avec une probabilité p ou l'échec avec une probabilité $q = 1-p$.
b) On répète n fois cette épreuve.
c) Les n épreuves sont indépendantes entre elles et la probabilité de réalisation de l'événement "succès" est la même à chaque épreuve et reste toujours égale à p.
Dans cette situation, on s'intéresse à la variable $X =$ "nombre de succès au cours des n épreuves".

4.2.2.2 Distribution de probabilités

Appelons X_i les variables de Bernoulli associées à chaque épreuve. Si la i-ème épreuve donne un succès, X_i vaut 1. Dans le cas contraire X_i vaut 0. La somme de ces variables comptabilise donc le nombre de succès au cours des n épreuves. On a donc $X = X_1 + X_2 + \cdots + X_n$. X peut prendre $n+1$ valeurs : $0, 1, \ldots, n$.
Cherchons la probabilité d'obtenir k succès, c'est-à-dire $\mathbb{P}(X=k)$, $k \in [\![0,n]\!]$.
La probabilité d'avoir k succès suivis de $n-k$ échecs est $p^k q^{n-k}$ car ces résultats sont indépendants les uns des autres.
La probabilité d'avoir k succès et $n-k$ échecs dans un autre ordre de réalisation est toujours $p^k q^{n-k}$. Donc tous les événements élémentaires qui composent l'événement $(X=k)$ ont une même probabilité.
Combien y en a-t-il ? Autant que de façons d'ordonner les k succès par rapport aux $n-k$ échecs. Il suffit de choisir les k places des succès parmi les n possibles et les $n-k$ échecs prendront les places restantes. Or il y a C_n^k manières de choisir k places parmi n.
Finalement, on obtient

$$\boxed{\mathbb{P}(X=k) = C_n^k p^k q^{n-k},\ k \in [\![0,n]\!].}$$

$\mathcal{B}(n,p)$

Définition 4.6 *On dit qu' une variable aléatoire X suit une loi* binomiale de paramètres n et p *si :*

$$\boxed{X(\Omega) = [\![0,n]\!],\ et\ \ \mathbb{P}(X=k) = C_n^k p^k q^{n-k},\ \forall k \in [\![0,n]\!].}$$

tirage — proba de succès

Notation 4.7 *On notera ceci en écrivant :* $\boxed{X \hookrightarrow \mathcal{B}(n,p).}$ Loi Binomiale

Remarque 4.3 *L'adjectif binomial vient du fait que lorsqu'on somme toutes ces probabilités, on retrouve le développement du binôme de Newton,*

$$\sum_{k=0}^{n} \mathbb{P}(X=k) = \sum_{k=0}^{n} C_n^k p^k q^{n-k} = (p+q)^n = 1.$$

Remarque 4.4 *La loi binomiale intervient aussi dans l'expérience de tirages avec remise.*
Dans une urne contenant N boules, dont R sont rouges et $N-R$ sont blanches$(1 \leqslant R \leqslant N-1)$.
Soit $p = \dfrac{R}{N}$ la probabilité de tirer une boule rouge. Si on fait n tirages successifs avec remise, alors si on note X le nombre de boules rouges obtenues au cours de ces n tirages, on a $X \hookrightarrow \mathcal{B}(n,p)$.(car on a n tirages indépendants dont les seules issues sont soit le succès "obtenir une boule rouge", soit l'échec "obtenir une boule blanche" avec la probabilité commune à tous les succès est la même égale à $p = \dfrac{R}{N}$).

4.2.2.3 Exemples

Exemple 4.4 *Sachant que la probabilité d'avoir un garçon est de 48 %, déterminez la probabilité pour une famille de cinq enfants d'avoir 3 garçons et 2 filles.*
Désignons par X la v.a. discrète qui donne le nombre de garçons dans une famille de cinq enfants. On a donc $X \hookrightarrow \mathcal{B}(5, 0.48)$. Donc :$\mathbb{P}(X = 3) = C_5^3 \times (0.48)^3 \times (1 - 0.48)^2$

Exemple 4.5 *Dans une entreprise, une machine produit des pièces dont les dimensions très précises doivent être respectées. Après un premier réglage, on constate qu'une proportion de 30 % de pièces sont défectueuses. On examine 5 pièces choisies au hasard dans la production.*
1. Quelle est la probabilité que 2 pièces soient défectueuses ?
2. Quelle est la probabilité qu'il n'y ai pas plus d'une pièce défectueuse ?
On appelle X le nombre de pièce défectueuses parmi les 5. On a donc : $X \hookrightarrow \mathcal{B}(5, 0.3)$. D'où
1. $\mathbb{P}(X = 2) = C_5^2 \times (0.3)^2 \times (1 - 0.3)^3$
2.$\mathbb{P}(X \leqslant 1) = \mathbb{P}(X = 0) + \mathbb{P}(X = 1) = C_5^0 \times (0.3)^0 \times (1 - 0.3)^5 + C_5^1 \times (0.3)^1 \times (1 - 0.3)^4$.

Proposition 4.8 (Somme de deux variables binomiales) *Si X_1 et X_2 sont des variables* indépendantes *qui suivent des lois binomiales $B(n_1, p)$ et $B(n_2, p)$ respectivement, alors $X_1 + X_2$ suit une loi binomiale $B(n_1 + n_2, p)$.*

4.2.3 Distribution géométrique

4.2.3.1 Situation concrète

a) On effectue une épreuve de Bernoulli. Elle n'a donc que deux issues : le succès avec une probabilité p ou l'échec avec une probabilité $q = 1 - p$.
b) On répète l'épreuve jusqu'à l'apparition du premier succès.
c) Toutes les épreuves sont indépendantes entre elles.

Dans cette situation, on s'intéresse à la variable X = "nombre de fois qu'il faut répéter l'épreuve pour obtenir le premier succès".

Remarque 4.5 *On est donc dans les mêmes hypothèses que pour la loi binomiale, mais le nombre d'épreuves n'est pas fixé à l'avance. On s'arrête au premier succès.*

4.2.3.2 Distribution de probabilités

L'ensemble des valeurs prises par X est $1, 2, 3, \ldots$. On cherche la probabilité d'avoir recours à k épreuves pour obtenir le premier succès.
Ce succès a une probabilité de réalisation de p. Puisque c'est le premier, il a été précédé de $k - 1$ échecs qui ont chacun eu la probabilité q de se produire. Étant donné l'indépendance des épreuves, on peut dire que la probabilité de réalisation de $k - 1$ échecs suivis d'un succès est le produit des probabilités de réalisation de chacun des résultats,

$$\boxed{\mathbb{P}(X = k) = q^{k-1}p}, \ k \in \mathbb{N}^*.$$

Définition 4.9 *On dit qu' une variable aléatoire X suit une loi* géométrique de paramètre p *si :*

$$\boxed{X(\Omega) = \mathbb{N}^*, \ \ et \ \ \mathbb{P}(X = k) = q^{k-1}p, \ \forall k \in \mathbb{N}^*.}$$

Notation 4.10 *On notera ceci en écrivant : $\boxed{X \hookrightarrow \mathcal{G}(p).}$*

Remarque 4.6 *L'appellation* géométrique *vient du fait qu'en sommant toutes les probabilités, on obtient une série géométrique. En effet,*

$$\sum_{k=1}^{+\infty} q^{k-1}p = \frac{p}{1-q} = 1.$$

4.2.3.3 Exemples

Exemple 4.6 *On lance une pièce de monnaie truquée dont la probabilité d'obtenir "pile" est p* $(p \neq \frac{1}{2})$*.*
On veut connaître la probabilité qu'il faille 7 jets pour obtenir "pile".
Soit X *le nombre de jets nécessaires pour obtenir "pile" pour la première fois. Donc* $X \hookrightarrow \mathcal{G}(p)$*. C'est à dire* $\mathbb{P}(X = k) = p \times (1-p)^{k-1}$*. D'où* $\mathbb{P}(X = 7) = (1-p)^6 \times p$*.*

Exemple 4.7 *Une urne contient 73 boules, 37 vertes et 36 jaunes. On répète l'expérience "tirer une boule avec remise". Quelle est la probabilité qu'il faille répéter 21 fois l'expérience pour obtenir la première boule verte ?*
Soit X *le nombre de tirages nécessaires pour obtenir la première boule verte. On voit que* $X \hookrightarrow \mathcal{G}(\frac{37}{73})$*.*
On a donc : $\mathbb{P}(X = 21) = \left(1 - \frac{37}{73}\right)^{20} \times \frac{37}{73}$*.*

4.2.4 Distribution de Poisson

La loi de Poisson est attribuée à Poisson, mathématicien français (1781-1840).

4.2.4.1 Situation concrète

Beaucoup de situations sont liées à l'étude de la réalisation d'un événement dans un intervalle de temps donné (arrivée de clients qui se présentent à un guichet d'une poste en une heure, apparitions de pannes d'un réseau informatique en une année, arrivée de malades aux urgences d'un hôpital en une nuit,....).
Les phénomènes ainsi étudiés sont des phénomènes d'attente ou de comptage.
On va voir que la loi de Poisson peut être interprétée comme un cas limite d'une loi binomiale.

Définition 4.11 *On peut considérer la loi de Poisson de paramètre* λ*,* $\lambda > 0$ *comme la loi limite d'une loi binomiale* $B(n, p_n)$ *lorsque le produit des paramètres* $n.p_n$ *converge vers* λ *quand* n *tend vers l'infini.*
On écrit $\boxed{X \hookrightarrow \mathbb{P}(\lambda).}$

Proposition 4.12 *La loi de Poisson de paramètre* λ *est donnée par*

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

$$\sum_{k=0}^{+\infty} \frac{\lambda^k}{k!} e^{-\lambda} = e^{-\lambda} \sum_{k=0}^{+\infty} \frac{\lambda^k}{k!}$$
$$= e^{-\lambda} . e^{\lambda}$$
$$= 1$$

Preuve. Si X_n suit une loi $B(n, p_n)$, on sait que pour k fixé et n assez grand tel que $n > k$ on a :

$$\begin{aligned}
\mathbb{P}(X_n = k) &= C_n^k (p_n)^k (1-p_n)^{n-k} \\
&= \left(\frac{np_n}{\lambda}\right)^k \frac{\lambda^k}{k!} \frac{n(n-1)\cdots(n-k+1)}{n^k}(1-p_n)^{n-k} \\
&= \left(\frac{np_n}{\lambda}\right)^k \frac{\lambda^k}{k!}\left[\frac{n}{n} \times \frac{n-1}{n} \times \cdots \times \frac{n-k+1}{n}\right](1-p_n)^{n-k}.
\end{aligned}$$

Le premier terme du membre de droite de la dernière expression converge vers 1 quand n tend vers l'infini. Chaque terme du produit entre crochets tend vers 1 lorsque n tend vers l'infini. Il y a k termes, c'est-à-dire un nombre fini. Donc le crochet tend vers 1. De plus,

$$\ln((1-p_n)^{n-k}) = (n-k)\ln(1-p_n) \sim n \times (-p_n) \sim -\lambda$$

donc $(1-p_n)^{n-k}$ tend vers $e^{-\lambda}$. On conclut que $\mathbb{P}(X_n = k)$ tend vers $e^{-\lambda}\lambda^k/k!$.

4.2.4.2 Exemples

Exemple 4.8 *Soit X le nombre d'appels téléphoniques reçus par un standard pendant un intervalle de temps l donné.*
On a donc $X \hookrightarrow \mathbb{P}(\lambda)$ où λ est le nombre d'appels moyen reçus par le standard pendant l'intervalle de temps l.

Exemple 4.9 *Soit X le nombre de véhicules se présentant à un péage d'autoroute pendant 1 heure.*
On a donc $X \hookrightarrow \mathbb{P}(\lambda)$ où λ est le nombre moyen de véhicules se présentant à un péage en 1 heure.

Exemple 4.10 *Un central téléphonique reçoit 100 appels par heure en moyenne.*
Calculez la probabilité que pendant deux minutes le central reçoive trois appels.
Soit X le nombre d'appels reçus en deux minutes. On a donc $X \hookrightarrow \mathbb{P}(\lambda)$ où λ est le nombre d'appels moyen reçus en deux minutes. On cherche λ, un produit en croix nous le donne :

$$\begin{array}{ccc} 1\ h = 60\ min & \longrightarrow & 100\ appels \\ 2\ min & \longrightarrow & \lambda \end{array}$$

Donc : $\lambda = \dfrac{2 \times 100}{60} = \dfrac{10}{3}$.
D'où : $\mathbb{P}(X = 3) = e^{-\frac{10}{3}} \times \dfrac{(\frac{10}{3})^3}{3!}$.

4.2.4.3 Approximation de la loi binomiale par la loi de Poisson

La loi binomiale dépend de deux paramètres n et p. Bien qu'il existe quelques tables, elle n'est pas simple à utiliser. La loi de Poisson ne dépend que d'un paramètre ce qui la rend plus pratique. Il faut donc avoir toujours présent à l'esprit que, lorsque les conditions le permettent, on peut avoir intérêt à remplacer une loi binomiale par une loi de Poisson.
Lorsque n est grand et p petit, de telle façon que le produit $np = \lambda$ reste petit par rapport à n, la loi binomiale $B(n, p)$ peut être approchée par la loi de Poisson $\mathbb{P}(\lambda)$ (revoir ce qui a été dit sur ce sujet dans le paragraphe "Distribution de probabilités"). Cette approximation s'appliquant lorsque p est petit, la loi de Poisson est appelée la loi des événements rares. En pratique, l'approximation est valable si $n > 20$, $p \leqslant 0.1$ et $np \leqslant 5$.

On approche la loi $B(n,p)$ par la loi $\mathbb{P}(np)$ dès que $n > 20$, $p \leqslant 0.1$ et $np \leqslant 5$.

Proposition 4.13 (Somme de deux lois de Poisson) *Si X_1 et X_2 sont des variables aléatoires* indépendantes *qui suivent des lois de Poisson de paramètres respectifs λ_1 et λ_2, alors $X_1 + X_2$ suit une loi de Poisson de paramètre $\lambda_1 + \lambda_2$.*

effectif total — effectif de succès — echantillon

4.2.5 Distribution Hypergéométrique $H(N, R, n)$

4.2.5.1 Situation concrète

Supposons qu'une urne contienne N boules, dont R sont rouges et $N - R$ sont blanches avec $1 \leqslant R \leqslant N - 1$.
Si on fait n tirages simultanés(donc sans remise), et si on note X le nombre de boules rouges obtenues au cours de ces n tirages, on a

$$\boxed{\mathbb{P}(X = k) = \frac{C_R^k \times C_{N-R}^{n-k}}{C_N^n}}, \text{ où } \underbrace{(0 \leqslant k \leqslant R)}_{①} \text{ et } \underbrace{(0 \leqslant n - k \leqslant N - R)}_{②} \tag{4.1}$$

On vérifie facilement que les conditions dans (4.1) se traduisent par

$$k \in [\![a, b]\!], \text{ où } a = \max(0, R + n - N) \text{ et } b = \min(R, n) \Leftrightarrow a \leqslant k \leqslant b$$

① $0 \leqslant k \leqslant R$

② $0 \leqslant n-k \leqslant N-R \Rightarrow R-N \leqslant k-n \leqslant 0 \Rightarrow \boxed{R-N+n \leqslant k \leqslant n} \Rightarrow$

Définition 4.14 *On dit qu' une variable aléatoire X suit une loi* hypergéométrique de paramètres N, *R et n si :*

$$\boxed{X(\Omega) = [\![a, b]\!], \text{ et } \mathbb{P}(X = k) = \frac{C_R^k \times C_{N-R}^{n-k}}{C_N^n}, \forall k \in [\![a, b]\!]}$$

où $a = \max(0, R + n - N),\ b = \min(R, n)$.

Notation 4.15 *On notera ceci en écrivant :* $X \hookrightarrow \mathcal{H}(N, R, n)$.

Remarque 4.7 - *Dans la notation ci-dessus, on a dans le cas général les significations suivantes :*

1. *N est l'effectif de la population totale ;*
2. *R est l'effectif de la sous-population à laquelle on s'intéresse ;*
3. *n est la taille de l'échantillon observé(ou tiré, étudié)*

Cette loi a été utilisée en cours et en T.D., elle est associée aux tirages simultanés sans remise, elle se généralise facilement pour plusieurs types(couleurs, genres, etc...)

Exemple 4.11 *Dans un jeu de 32 cartes, on tire une main de 5 cartes. Soit X le nombre de cœurs obtenus. Donner la loi de X.*
X est à valeurs dans $E = [\![0, 5]\!]$. On cherche $\mathbb{P}(X = k), k \in [\![0, 5]\!]$.
Justifier que : $\mathbb{P}(X = k) = \dfrac{C_8^k \times C_{24}^{5-k}}{C_{32}^5}, \forall k \in [\![0, 5]\!]$

Proposition 4.16 (voir série 1) *On a la formule de* *<u>Vandermonde</u>* *suivante :*

$$\sum_{i+j=k} C_n^i C_m^j \quad = \quad C_{n+m}^k \quad ,$$

où i, j, k, n, m sont des entiers naturels tels que $0 \leqslant k \leqslant n + m$.

Pour montrer cette formule, on pourra utiliser le produit de polynômes.

Remarque 4.8 *La proposition ci-dessus, appelée "Formule de Vandermonde", permet de remarquer que*

$$\sum_{k=a}^{b} \mathbb{P}(X = k) = \sum_{k=a}^{b} \frac{C_R^k \times C_{N-R}^{n-k}}{C_N^n} = \frac{\sum_{k=a}^{b} C_R^k \times C_{N-R}^{n-k}}{C_N^n} = \frac{C_N^n}{C_N^n} = 1$$

4.2.6 Distribution Uniforme $\quad U(\{x_1, \cdots, x_m\})$

Définition 4.17 *On dit qu' une variable aléatoire X suit une loi* uniforme *sur l'ensemble $\{x_1, x_2, \cdots, x_n\}$ si :*

$$\boxed{X(\Omega) = \{x_1, x_2, \cdots, x_n\} \text{ et } \mathbb{P}(X = x_k) = \frac{1}{n}, \forall k \in [\![1, n]\!].}$$

Notation 4.18 *On notera ceci en écrivant :* $X \hookrightarrow \mathcal{U}\{x_1, x_2, \cdots, x_n\}$.

Exemple 4.12 *On lance un dé non pipé. Soit X le nombre de points de la face supérieure.*
Donner la loi de X puis calculer les probabilités suivantes : $\mathbb{P}(X \leqslant 3)$, $\mathbb{P}(X > 1)$ et $\mathbb{P}(1 < X \leqslant 3)$.
X est à valeurs dans $E = [\![1, 6]\!]$. On cherche $\mathbb{P}(X = k), k \in [\![1, 6]\!]$.
On a $X \hookrightarrow \mathcal{U}\{1, 2, 3, 4, 5, 6\}$ et donc $\mathbb{P}(X = k) = 1/6, \forall k \in [\![1, 6]\!]$.
On pourra facilement calculer les probabilités demandées. On aura ainsi :

- $\mathbb{P}(X \leqslant 3) = \mathbb{P}(X = 1) + \mathbb{P}(X = 2) + \mathbb{P}(X = 3) = \dfrac{3}{6}$.
- $\mathbb{P}(X > 1) = 1 - \mathbb{P}(X = 1) = 1 - \dfrac{1}{6} = \dfrac{5}{6}$.
- $\mathbb{P}(1 < X \leqslant 3) = \mathbb{P}(X = 2) + \mathbb{P}(X = 3) = \dfrac{2}{6}$.

8

4.3 Fonction de répartition $F_x(t) = P(X \leq t)$

Définition 4.19 (Fonction de répartition) *Soit X une variable aléatoire discrète. On appelle fonction de répartition de X la fonction notée :*

$$\begin{aligned} F_X : \mathbb{R} &\to [0,1] \\ x &\mapsto F_X(x) = \mathbb{P}(X \leqslant x) \end{aligned}$$

Exemple 4.13 *Soit X une v.a.d. à valeur dans $\mathbb{N}$. Quelle est la valeur de $F_X(2.3)$?*
$F_X(2.3) = \mathbb{P}(X \leqslant 2.3) = \mathbb{P}(X=0) + \mathbb{P}(X=1) + \mathbb{P}(X=2)$.

Exemple 4.14
En reprenant l'exercice où l'on tire une main de 5 cartes dans un jeu de 32, avec X le nombre de cœurs. Donner $F_X(x), \forall x \in \mathbb{R}$.

X est à valeur dans $[\![0,5]\!]$ et $\mathbb{P}(X=k) = \dfrac{C_8^k \times C_{24}^{5-k}}{C_{32}^5}$.
De plus on a : $F_X(x) = \mathbb{P}(X \leqslant x)$

- *1^er^cas, si $x < 0$, alors : $F_X(x) = \mathbb{P}(\emptyset) = 0$;*
- *2^e^cas, si $0 \leqslant x < 1$, alors : $F_X(x) = \mathbb{P}(X \leqslant x) = \mathbb{P}(X=0)$;*
- *3^e^cas, si $1 \leqslant x < 2$, alors : $F_X(x) = \mathbb{P}(X \leqslant x) = \mathbb{P}(X=0) + \mathbb{P}(X=1)$;*

  $\vdots$
- *6^e^cas, si $4 \leqslant x < 5$, alors :*
  $F_X(x) = \mathbb{P}(X \leqslant x) = \mathbb{P}(X=0) + \mathbb{P}(X=1) + \mathbb{P}(X=2) + \mathbb{P}(X=3) + \mathbb{P}(X=4)$;
- *7^e^cas, si $x \geqslant 5$, alors :*
  $F_X(x) = \mathbb{P}(X \leqslant x) = \mathbb{P}(X=0) + \mathbb{P}(X=1) + \mathbb{P}(X=2) + \mathbb{P}(X=3) + \mathbb{P}(X=4) + \mathbb{P}(X=5) = 1$.

Dessinons le graphe de F_X.

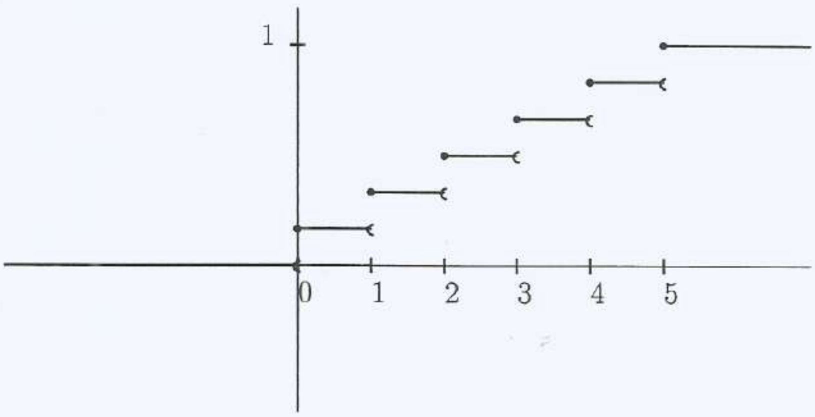

FIGURE 4.1 – Courbe d'une fonction de répartition dans le cas d'une v.a. discrète

4.3.0.1 Intérêt de la fonction de répartition

Si on connaît F_X on est capable de calculer :

$a < b \Rightarrow [X \leq a] \subset [X \leq b]$

1. $\{\mathbb{P}(X=k); k \in E\}$, c'est à dire la loi de X,
2. $\mathbb{P}(a < X \leqslant b)$.

En effet :

$P(a < X \leq b) = P\left([X \leq b] \cap \overline{[X \leq a]}\right) = P([X \leq b]) - P([X \leq a])$
$= F_x(b) - F_x(a)$

9

1. On a : $(X \leqslant k) = (X = k) \cup (X \leqslant k-1)$. Ces deux événements sont incompatibles, d'où :

$$\underbrace{\mathbb{P}(X \leqslant k)}_{F_X(k)} = \mathbb{P}(X = k) + \underbrace{\mathbb{P}(X \leqslant k-1)}_{F_X(k-1)}$$

On en conclut que : $\mathbb{P}(X = k) = F_X(k) - F_X(k-1)$

2. On a : $(X \leqslant b) = (X \leqslant a) \cup (a < X \leqslant b)$. Ces événements sont incompatibles, donc :

$$\underbrace{\mathbb{P}(X \leqslant b)}_{F_X(b)} = \underbrace{\mathbb{P}(X \leqslant a)}_{F_X(a)} + \mathbb{P}(a < X \leqslant b)$$

D'où : $\mathbb{P}(a < X \leqslant b) = F_X(b) - F_X(a)$.

4.4 Paramètres d'une loi de probabilité discrète

4.4.1 Espérance

Définition 4.20 (Espérance) *Soit X une variable aléatoire discrète à valeurs dans $E (E \subset \mathbb{N})$. On appelle espérance de X le nombre réel (s'il existe) noté $E(X)$ défini par :*

$$\boxed{E(X) = \sum_{k \in E} k \times \mathbb{P}(X = k)}$$

Exemple 4.15 *On lance deux fois de suite un dé. On pose X le nombre de 6 obtenus. On cherche $E(X)$.*
X est à valeurs dans $\{0, 1, 2\} = [\![0, 2]\!]$.

$$\begin{aligned} E(X) &= \sum_{k=2}^{2} k \times \mathbb{P}(X = k) = 0 \times \mathbb{P}(X = 0) + 1 \times \mathbb{P}(X = 1) + 2 \times \mathbb{P}(X = 2) \\ &= 0 + 1 \times \frac{10}{36} + 2 \times \frac{1}{36} = \frac{12}{36} = \frac{1}{3} \end{aligned}$$

Définition 4.21 (Espérance de $f(X)$) *Pour toute fonction réelle f, on définira l'espérance de $f(X)$ notée par $E(f(X))$ lorsqu'elle existe par :*

$$\boxed{E(f(X)) = \sum_{k \in E} f(k) \times \mathbb{P}(X = k)}$$

Exemple 4.16 *En reprenant le précédant exemple :*

$$\begin{aligned} \boxed{E(X^2) = \sum_{k=2}^{2} k^2 \times \mathbb{P}(X = k)} \\ = 0 \times \mathbb{P}(X = 0) + 1 \times \mathbb{P}(X = 1) + 4 \times \mathbb{P}(X = 2) = \frac{14}{36} \end{aligned}$$

Proposition 4.22 *Soit $(a, b) \in \mathbb{R}^2$, on a : $\boxed{E(aX + b) = aE(X) + b.}$*

Preuve.

$$\begin{aligned} E(aX + b) &= \sum_{k \in E} (ak + b) \times \mathbb{P}(X = k) = \sum_{k \in E} ak\mathbb{P}(X = k) + \sum_{k \in E} b\mathbb{P}(X = k) \\ &= a \sum_{k \in E} k\mathbb{P}(X = k) + b \sum_{k \in E} \mathbb{P}(X = k) = aE(X) + b \end{aligned}$$

Définition 4.23 (Variable aléatoire centrée) *On dit que X (variable aléatoire discrète) est centrée si son espérance est nulle.*

Remarque 4.9 *Si X n'est pas centrée, $X - E(X)$ est centrée.*
En effet $E(X - E(X)) = E(X) - E(X) = 0$ (car $E(X)$ est une constante).

Proposition 4.24 (Linéarité de l'espérance) *Soit $(a, b) \in \mathbb{R}^2$, on a :*

$$E(aX + bY) = aE(X) + bE(Y)$$

Preuve. (admise)

4.4.2 Variance-Écart-type

Définition 4.25 (Variance) *Soit X une variable aléatoire discrète, on appelle variance de X, le nombre réel (s'il existe) noté $V(X)$ défini par :*

$$V(X) = E\left((X - E(X))^2\right)$$

Définition 4.26 (Écart-type) *On appelle écart-type de X le nombre réel (s'il existe) noté $\sigma(X)$ défini par :*

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

Ceci mesure la dispersion de X par rapport à son espérance.

Proposition 4.27 *On a :*

1. $V(X) = E\left(X^2\right) - (E(X))^2$
2. $V(X) = E(X(X-1)) + E(X) - (E(X))^2$

Preuve.

1.

$$\begin{aligned} V(X) &= E\left((X - E(X))^2\right) = E\left(X^2 - 2XE(X) + (E(X))^2\right) \\ &= E\left(X^2\right) - 2E(X)E(X) + \left(E(X)^2\right) = E\left(X^2\right) - 2(E(X))^2 + \left(E(X)^2\right) \\ &= E\left(X^2\right) - (E(X))^2 \end{aligned}$$

2.

$$\begin{aligned} & E(X(X-1)) + E(X) - (E(X))^2 = E\left(X^2 - X\right) + E(X) - (E(X))^2 \\ = \; & E\left(X^2\right) - E(X) + E(X) - (E(X))^2 = E\left(X^2 - 2E(X) + (E(X))^2\right) = V(X) \end{aligned}$$

Proposition 4.28

$$V(a \times X + b) = a^2 \times V(X)$$

Preuve.

$$\begin{aligned} V(a \times X + b) &= E((a \times X + b)^2) - (E(a \times X + b))^2 \\ &= E(a^2 \times X^2 + 2ab \times X + b^2) - (a \times E(X) + b)^2 \\ &= a^2 \times E(X^2) + 2ab \times E(X) + b^2 - (a^2 \times (E(X))^2 + 2ab \times E(X) + b^2) \\ &= a^2 \times (E(X^2) - (E(X))^2 \\ &= a^2 \times V(X) \end{aligned}$$

4.4.3 Espérance et variance des lois binomiale, de Poisson, géométrique et Uniforme

Théorème 4.29
Si $X \hookrightarrow B(n,p)$, *alors* $E(X) = n \times p$ *et* $V(X) = n \times p \times (1-p) = npq$.
Si $X \hookrightarrow \mathbb{P}(\lambda)$, *alors* $E(X) = \lambda$ *et* $V(X) = \lambda$.
Si $X \hookrightarrow G(p)$, *alors* $E(X) = \dfrac{1}{p}$ *et* $V(X) = \dfrac{1-p}{p^2} = \dfrac{q}{p^2}$.
Si $X \hookrightarrow \mathcal{U}\{x_1, x_2, \cdots, x_n\}$ *alors* $E(X) = \overline{x}$ *et* $V(X) = V(x)$ *où* (x) *est la série* $(x_1, x_2, \cdots, x_n)$.

Preuve des résultats pour la loi binomiale :

— Pour l'espérance :

$$\begin{aligned}
\boxed{E(X)} &= \sum_{k=0}^{n} k \times \mathbb{P}(X=k) = \sum_{k=0}^{n} k \times C_n^k \times p^k \times (1-p)^{n-k} \\
&= \sum_{k=1}^{n} k \times \frac{n!}{k! \times (n-k)!} \times p^k \times (1-p)^{n-k} = \sum_{k=1}^{n} \frac{n!}{(k-1)! \times (n-k)!} \times p^k \times (1-p)^{n-k} \\
&= n \times p \times \sum_{k=1}^{n} \frac{(n-1)!}{(k-1)! \times (n-k)!} \times p^{k-1} \times (1-p)^{n-k} \\
&= n \times p \times \underbrace{\sum_{k=1}^{n} \times C_{n-1}^{k-1} \times p^{k-1} \times (1-p)^{n-k}}_{(p+(1-p))^{n-1}} \\
&= n \times p \times 1^{n-1} = \boxed{n \times p}
\end{aligned}$$

$V(x) = E(X^2) + (E(x))^2$
$= E(x(x-1)+x) + (E(x))^2$
$= E(x(x-1)) + E(X) + (E(X))^2$

— Pour la variance :

$$\boxed{V(X) = E(X(X-1)) + E(X) - (E(X))^2}$$

$$\begin{aligned}
\boxed{E(X(X-1))} &= \sum_{k=0}^{n} k \times (k-1) \times \mathbb{P}(X=k) = \sum_{k=2}^{n} k \times (k-1) \times C_n^k \times p^k \times (1-p)^{n-k} \\
&= \sum_{k=2}^{n} k \times (k-1) \times \frac{n!}{k! \times (n-k)!} \times p^k \times (1-p)^{n-k} \\
&= \sum_{k=2}^{n} \frac{n!}{(k-2)! \times (n-k)!} \times p^k \times (1-p)^{n-k} \\
&= n \times (n-1) \times p^2 \sum_{k=2}^{n} \times \frac{(n-2)!}{(k-2)! \times (n-k)!} \times p^{k-2} \times (1-p)^{n-k} \\
&= n \times (n-1) \times p^2 \underbrace{\sum_{k=2}^{n} \times C_{n-2}^{k-2} \times p^{k-2} \times (1-p)^{n-k}}_{(p+(1-p))^{n-2}} \\
&= n \times (n-1) \times p^2 \times 1^{n-2} = n \times (n-1) \times p^2 \\
\text{d'où } V(X) &= n \times (n-1) \times p^2 + n \times p - n^2 \times p^2 \\
&= n \times p[(n-1) \times p + 1 - n \times p] \\
&= n \times p \times (1-p) = \boxed{npq}
\end{aligned}$$

Pour la preuve des autres résultats voir le cours.

Exercice résolu 1 : *Loi exacte et approximation*

2 % des articles produits par une usine sont défectueux.
On note X la v.a. discrète désignant le nombre de produits défectueux sur un lot de 200 articles.

1)

a) Donner, en la justifiant, la loi de la v.a. X.

b) Calculer directement $P(X=4)$(à 10^{-4} près).

2)

a) Par quelle loi peut-on approcher la loi de X ?

b) Donner, en utilisant l'approximation ci-dessus, une valeur approchée de $P(X=4)$(à 10^{-4} près).

c) Comparer les résultats trouvés puis conclure.

Solution

1)

a) On a un lot de $n=200$ produits choisis au hasard(donc implicitement on a effectué des choix indépendants). Chaque produit choisi a deux possibilités : il est soit défectueux(ici ceci représente un succès), soit non défectueux(échéc). Le texte suggère que la probabilité de trouver un article choisi défectueux est commune à tous les articles et vaut $p=0.02$.

Le nombre d'articles défectueux représente le nombre de succès sur les 200 articles produits, ainsi X suit la loi binomiale de paramètres $n=200$ et $p=0.02$ qu'on note par $X \hookrightarrow \mathcal{B}(n=200, p=0.02)$, d'où

$$X(\Omega)=[\![0,200]\!], \text{ et } P(X=k)=C_{200}^{k} \quad 0.02^{k}0.98^{200-k}, \ \forall k \in [\![0,200]\!].$$

b) Pour calculer $P(X=4)$(à 10^{-4} près), il suffit d'utiliser la dernière formule avec $k=4$, ainsi on a :

$$P(X=4)=C_{200}^{4} \quad 0.02^{4}0.98^{196}=0.1973=v_{exacte}$$

2)

a) La loi binomiale $\mathcal{B}(n=200, p=0.02)$ (car $n=200>20, p=0.02 \leqslant 0.1$ et $n \times p=4 \leqslant 5$)peut être approchée par la loi de poisson $\mathcal{P}(\lambda=n \times p=4)$. Ainsi

$$P(X=k) \cong e^{-\lambda}\frac{\lambda^{k}}{k!}, \ k \in [\![0,200]\!].$$

b) Une valeur approximative de $P(X=4)$(à 10^{-4} près) utilisant la loi de poisson de paramètre 4 est donc obtenue par la dernière formule et vaut

$$P(X=4) \cong e^{-4}\frac{4^{4}}{4!}=0.1954=v_{approximative}.$$

c) Pour comparer deux résultats il est recommandé de calculer l'erreur relative qu'on calcule par la formule suivante :

$$\text{Erreur relative}=\frac{|\text{valeur approchée -valeur exacte}|}{|\text{valeur exacte}|}=\frac{|0.1954-0.1973|}{|0.1973|}=0.0101.$$

En multipliant par 100 le dernier résultat on trouve un pourcentage d'erreur relative 1.01% qui est "acceptable". Cette erreur relative est faible car n est très grand et p très petit et np relativement petit : $n=200>20, p=0.02 \leqslant 0.1$ et $n \times p=4 \leqslant 5$. Pour d'autres applications, en ayant par exemple $np=10$, on aurait une erreur relative très grande donc l'approximation serait mauvaise !

Exercice résolu 2 : *Somme de deux v.a.d indépendantes*

Deux v.a. X et Y sont indépendantes ssi $P([X \in A] \cap [Y \in B]) = P[X \in A] \times P[Y \in B]$, $\forall A, B$.

1) Soient X_1 et X_2 deux v.a. indépendantes telles que : $X_i \hookrightarrow \mathcal{B}(n_i, p)$ pour $i = 1, 2$ où les $n_i \in \mathbb{N}^*$ et $0 < p < 1$. Montrer $X_1 + X_2 \hookrightarrow \mathcal{B}(n_1 + n_2, p)$.

2) Soient X_1 et X_2 deux v.a. indépendantes telles que : $X_i \hookrightarrow P(\lambda_i)$ pour $i = 1, 2$ où les $\lambda_i > 0$. Montrer que $X_1 + X_2 \hookrightarrow P(\lambda_1 + \lambda_2)$.

Solution

Posons $Z = X_1 + X_2$.

1) $X_i(\Omega) = [\![0, n_i]\!]$ pour $i = 1, 2$ donc $Z(\Omega) \subset [\![0, n_1 + n_2]\!]$.

$$\begin{aligned}
P\{Z = k\} &= \sum_{l=0}^{k} P\{X_1 = l \text{ et } X_2 = (k-l)\} \\
&= \sum_{l=0}^{k} P\{X_1 = l\} P\{X_2 = (k-l)\} \\
&= \sum_{l=0}^{n_1} C_{n_1}^{l} p^l q^{n_1-l} \times C_{n_2}^{k-l} p^{k-l} q^{n_2-k+l} \\
&= p^k q^{n_1+n_2-k} \sum_{l=0}^{n_1} C_{n_1}^{l} \times C_{n_2}^{k-l} (\text{ et via la formule de Vandermonde, on aura }) \\
&= C_{n_1+n_2}^{k} p^k q^{n_1+n_2-k} \neq 0
\end{aligned}$$

D'où $Z(\Omega) = [\![0, n_1 + n_2]\!]$ et Z est encore binomiale, de paramètres $(n_1 + n_2, p)$

2) $X_i(\Omega) = \mathbb{N}$ pour $i = 1, 2$ donc $Z(\Omega) \subset \mathbb{N}$.

Pour k entier $\geqslant 0$, en utilisant l'indépendance et la définition, on a :

$$\begin{aligned}
P(X_1 + X_2 = k) &= P([X_1 = 0] \cap [X_2 = k]) \cup ([X_1 = 1] \cap [X_2 = k-1]) \cup \cdots \cup ([X_1 = k] \cap [X_2 = 0])) \\
&= \sum_{i=0}^{k} P([X_1 = i] \cap [X_2 = k-i]) = \sum_{i=0}^{k} P[X_1 = i] \times P[X_2 = k-i] \\
&= \sum_{i=0}^{k} e^{-\lambda_1} e^{-\lambda_2} \times \frac{\lambda_1^i \lambda_2^{k-i}}{i! \times (k-i)!} \\
&= e^{-(\lambda_1+\lambda_2)} \times \sum_{i=0}^{k} \frac{\lambda_1^i \lambda_2^{k-i}}{i! \times (k-i)!} \\
&= \frac{e^{-(\lambda_1+\lambda_2)}}{k!} \times \sum_{i=0}^{k} \frac{k!}{i! \times (k-i)!} \lambda_1^i \lambda_2^{k-i} \\
&= \frac{e^{-(\lambda_1+\lambda_2)}}{k!} \times \sum_{i=0}^{k} C_k^i \lambda_1^i \lambda_2^{k-i} \\
&= \frac{e^{-(\lambda_1+\lambda_2)}}{k!} \times (\lambda_1 + \lambda_2)^k \neq 0
\end{aligned}$$

D'où $Z(\Omega) = \mathbb{N}$ et Z est encore poissonienne, de paramètre $(\lambda_1 + \lambda_2)$.

Résumé des lois usuelles discrètes

Nom et paramètres	$X(\Omega)$	$p_k = P[X=k], k \in X(\Omega)$	$E(X)$	$V(X)$
Bernoulli $\mathcal{B}(p), p \in]0,1[$	$\{0,1\}$	$p_1 = p, p_0 = 1-p = q$	p	pq
Binomiale $\mathcal{B}(n,p), n \in \mathbb{N}^*, p \in]0,1[$	$[\![0,n]\!]$	$p_k = C_n^k p^k q^{n-k}$	np	npq
Poisson $\mathcal{P}(\lambda), \lambda > 0$	$\mathbb{N}$	$p_k = \frac{\lambda^k}{k!} e^{-\lambda}$	λ	λ
Géométrique $G(p), p \in]0,1[$	$\mathbb{N}^*$	$p_k = q^{k-1} p$	$\frac{1}{p}$	$\frac{q}{p^2}$
Hypergéométrique $\mathcal{H}(N,R,n)(*1)$	$[\![a,b]\!](*2)$	$p_k = \frac{C_R^k C_{N-R}^{n-k}}{C_N^n}$	$(*3)$	$(*4)$
Uniforme sur $\{x_1, \cdots, x_n\}, n \in \mathbb{N}^*$	$\{x_1, \cdots, x_n\}$	$p_k = P[X = x_k] = \frac{1}{n}$	$\overline{x}$	$V(x)$

$(*1)$: $(N,R,n) \in (\mathbb{N}^*)^3, \max(n,R) < N$, $(*2)$: $a = \max(0, R+n-N)$, $b = \min(R,n)$.
$(*3)$ et $(*4)$: la détermination de ces paramètres est hors programme.

Conditions d'utilisation

1. $X \hookrightarrow \mathcal{B}(p)$ la loi de **Bernoulli** de paramètre p si X prend deux valeurs possibles 1 et 0. $[X=1]$ désignera le succès et $[X=0]$ désignera l'échec. La loi peut être associée à l'expérience "succès-échec". $\boldsymbol{p}$ désignera la probabilité du succès.
2. $X \hookrightarrow \mathcal{B}(n,p)$ la loi **binomiale** de paramètres n et p si X désigne le **nombre** de succès au cours de $\boldsymbol{n}$ épreuves **indépendantes** dont chacune a **deux** issues possibles ; succès et échec, avec une probabilité **commune** pour tous les succès égale à $\boldsymbol{p}$. La loi peut être associée aussi aux tirages **avec remise** dans une population constituée de deux catégories d'individus. Cette loi se généralise en multinomiale lorsqu'il y a plusieurs catégories.
3. $X \hookrightarrow \mathcal{P}(\lambda)$ la loi de **Poisson** de paramètre λ si X est associée aux phénomènes de **comptage** ou **d'attente** ou apparaît comme loi **limite** $\mathcal{B}(n, p_n)$ lorsque np_n tend vers λ quand n tend vers $+\infty$.
4. $X \hookrightarrow G(p)$ la loi de **géométrique** de paramètre p si X est la loi du **premier** succès associée à des épreuves **indépendantes** à **deux** issues : succès et échec, avec une probabilité **commune** pour tous les succès égale à $\boldsymbol{p}$.
5. $X \hookrightarrow \mathcal{H}(N,R,n)$ la loi **Hypergéométrique** de paramètres $\boldsymbol{N, R, n}$ si X désigne le nombre d'éléments de la sous-population à laquelle on s'interesse d'effectif $\boldsymbol{R}$ dans une population mère d'effectif $\boldsymbol{N}$ lorsqu'on prélève **sans remise** $\boldsymbol{n}$ éléments. Cette loi se généralise pour plusieurs sous-populations.
6. $X \hookrightarrow \mathcal{U}(x_1, \cdots, x_n)$ la loi **uniforme** sur $\{x_1, \cdots, x_n\}$ si X peut prendre chacune des $\boldsymbol{n}$ valeurs x_i avec la même probabilité $\frac{1}{n}$.