Annexe 2 : Rappels sur les variables aléatoires discrètes.

1. Définitions et vocabulaire

Rappel : Soit (Ω, T, P) un *espace probabilisé*, c'est-à-dire $\begin{cases} \Omega = \text{Univers (possibilités)} \\ T = \text{Tribu (événements)} \\ P = \text{Probabilité} \end{cases}$

Une *variable aléatoire* X est une application de Ω dans ℝ telle que,

pour tout nombre *x*, $\{\omega \in \Omega, X(\omega) \leq x\} \in T$.

Remarques : En pratique, nous n'aurons jamais accès ni à l'univers Ω ni à la tribu *T*.

On retiendra simplement qu'une variable aléatoire est une « valeur aléatoire ».

Exemple : X = Taille d'un individu choisi au hasard. $\begin{cases} \Omega = \text{Tous les humains} \\ T = \text{Groupes d'humains} \\ P = \text{Probabilité de choisir tel ou tel groupe} \end{cases}$

Définitions :

Le *support d'une variable aléatoire*, noté $X(\Omega)$, est l'ensemble des valeurs prises par X.

Une variable aléatoire est *discrète* si on peut « numéroter » son support :

$$X(\Omega) = \{x_0, x_1, x_2, x_3, \dots\}$$

Exemples :

- Pour X_1 = résultat d'un dé à 4 faces , on a $X_1(\Omega) =$ **{ 1 , 2 , 3 , 4 } = [[1 , 4]]**
- Pour X_2 = âge, en années, d'une personne prise au hasard, on a $X_2(\Omega) =$ **[[0 , 120]]**
- Pour X_3 = gains algébriques, en euros, à une partie de poker , on a $X_3(\Omega) =$ $\mathbb{Z}$
- Pour X_4 = longueur d'un saut en longueur , on a $X_4(\Omega) =$ **[0 , 10]**

Parmi ces exemples, les variables aléatoires **X_1 , X_2 et X_3** sont discrètes.

Remarque : Typiquement, on ne peut pas numéroter les valeurs d'une intervalle réel.

Définition : La *loi d'une variable aléatoire discrète* est la donnée de :

- son support $X(\Omega)$
- l'ensemble des probabilités $P(X=x)$ où $x \in X(\Omega)$.

Remarque : Si possible, on présente cette loi sous forme de tableau.

Exercice : Une urne contient 1 boule bleue et 3 boules jaunes. On tire les boules au hasard jusqu'à ce qu'il ne reste plus que des boules d'une seule couleur dans l'urne.

Soit X le nombre de tirages nécessaires.

1. Quel est l'ensemble X(Ω) des valeurs prises par X ?

X(Ω) = { 1 , 2 , 3 }

2. Déterminer la loi de X. *On pourra s'aider d'un arbre...*

Valeur k	**1**	**2**	**3**
P(X=k)	**1 / 4**	**3/4 * 1/3 = 1/4**	**1/2**

Proposition : Une suite de valeur $(p_0, p_1, p_2, \ldots)$ définit une loi de probabilités ssi :

- $\forall n \in \mathbb{N},\ p_n \geq 0$ *Chaque valeur est positive.*
- $\sum_{n=0}^{+\infty} p_n = 1$ La somme/série des valeurs vaut 1

Exemples :

- Dans l'exercice précédent,

  toutes les valeurs de la 2e ligne sont positives et leur somme vaut 4/4 = 1

- On considère $p_n = \frac{1}{n} - \frac{1}{n+1}$ pour $n \geq 1$

  D'une part $p_n \geq 0$ puisque $\frac{1}{n} \geq \frac{1}{n+1}$

  D'autre part $\sum_{n=1}^{N} p_n = \sum_{n=1}^{N} \left(\frac{1}{n} - \frac{1}{n+1}\right) = 1 - \frac{1}{N+1}$ par télescopage. Avec la limite $N \to +\infty$ on conclut que $\sum_{n=0}^{+\infty} p_n = \lim_{N \to +\infty} \left(1 - \frac{1}{N+1}\right) = 1$ i.e (p_n) définit une loi de probabilités.

Définition : Soit X une variable aléatoire discrète.

La *fonction de répartition de* X est la fonction définie par $\begin{array}{llll} F_X : & \mathbb{R} & \to & [0,1] \\ & x & \to & P(X \leq x) \end{array}$

La connaissance de la fonction de répartition F_X équivaut à la connaissance de la loi de X.

Exercice : On lance un dé équilibré à 4 faces à trois reprises. On note X_1, X_2, X_3 les résultats obtenus et X le plus grand de ces résultats.

1. Quel est le support de X ? **X(Ω) = { 1 , 2 , 3 , 4 }**

2. Calculer $P(X=1)$ en détaillant les événements utilisés.

$P(X=1)=P(X_1=1 \cap X_2=1 \cap X_3=1)=P(X_1=1)\times P(X_2=1)\times P(X_3=1)$ par indépendance

donc $P(X=1)=\frac{1}{4}\times\frac{1}{4}\times\frac{1}{4}=\frac{1}{64}$

3. Justifier que $(X\leq 2)=(X_1\leq 2)\cap(X_2\leq 2)\cap(X_3\leq 2)$ et en déduire $P(X\leq 2)$.

Le maximum des 3 dés est inférieur à 2 ssi tous les dés ont une valeur inférieure ou égale à 2.

Ainsi $(X\leq 2)=(X_1\leq 2)\cap(X_2\leq 2)\cap(X_3\leq 2)$ **puis**

$$P(X\leq 2)=P(X_1\leq 2)\times P(X_2\leq 2)\times P(X_3\leq 2)=\frac{2}{4}\times\frac{2}{4}\times\frac{2}{4}=\frac{8}{64}$$ **par indépendance.**

4. En déduire $P(X=2)$.

$$P(X=2)=P(X\leq 2)-P(X=1)=\frac{8}{64}-\frac{1}{64}=\frac{7}{64}$$

5. Déterminer la loi de X, en vous inspirant de cette méthode.

$$(X\leq 3)=(X_1\leq 3)\cap(X_1\leq 3)\cap(X_1\leq 3)$$

donc $P(X\leq 3)=P(X_1\leq 3)\times P(X_2\leq 3)\times P(X_3\leq 3)=\frac{3}{4}\times\frac{3}{4}\times\frac{3}{4}=\frac{27}{64}$ **par indépendance.**

On conclut que $P(X=3)=P(X\leq 3)\ -\ P(X\leq 2)=\frac{27}{64}-\frac{8}{64}=\frac{19}{64}$

Pour finir, on calcule $P(X=4)=1-P(X\leq 3)=\frac{37}{64}$

Conclusion : la loi de X est

Valeur k	**1**	**2**	**3**	**4**
Probabilité P(X=k)	**1/64**	**7/64**	**19/64**	**37/64**

Remarque : De façon générale, si X prend des valeur entières, on peut écrire

$$P(X=k)=P(X\leq k)-P(X\leq k-1)$$

Cette égalité permet d'obtenir la loi de X à partir de sa fonction de répartition.

2. Espérance et variance d'une variable aléatoire discrète.

<u>**Définition :**</u> Soit X une variable aléatoire discrète de support $X(\Omega)=\{x_0, x_1, x_2, x_3, \ldots\}$.

On dit que X *admet une espérance* si la série $\sum_k x_k P(X=x_k)$ est <u>absolument convergente</u>.

Dans ce cas, *l'espérance de* X est définie par $E(X)=\sum_k x_k P(X=x_k)$.

Interprétation : L'espérance de X mesure la « *moyenne* » des valeurs prises par X **sur un grand nombre de répétitions de X réalisées indépendamment les unes des autres.**

Cas simples :

- Si le support de X est fini, la série est en fait une somme finie.
  Dans ce cas, X admet automatiquement une espérance.
- Si X ne prend que des valeurs positives, on peut confondre *convergence* et *convergence absolue* de cette série.

Exemples :

1. On reprend l'exemple précédent. X = maximum de 3 dés à 4 faces.

La loi de X est

Valeur k	**1**	**2**	**3**	**4**
Probabilité P(X=k)	$\frac{1}{64}$	$\frac{7}{64}$	$\frac{19}{64}$	$\frac{37}{64}$

On en déduit $E(X)=1\times\frac{1}{64}+2\times\frac{7}{64}+3\times\frac{19}{64}+4\times\frac{37}{64}=\frac{220}{64}=\frac{55}{16}$

2. On considère la variable aléatoire Y définie sur $\mathbb{N}^*$ par $P(Y=n)=\frac{1}{n}-\frac{1}{n+1}$.

Nous avons vu (Page 2) que ceci définit bien une loi de probabilités : valeurs positives de somme 1.

Pour l'espérance de Y, étudions

$$\sum_{n=1}^{N}|n P(Y=n)|=\sum_{n=1}^{N} n\left(\frac{1}{n}-\frac{1}{n+1}\right)=\sum_{n=1}^{N}\frac{n}{n(n+1)}=\sum_{n=1}^{N}\frac{1}{n+1}$$

On peut prouver (Cf. TD1, Exercice 8) que la série $\sum_{n=1}^{+\infty}\frac{1}{n+1}$ diverge.

Par conséquent la variable aléatoire Y n'admet pas d'espérance.

<u>Théorème de transfert :</u> Soit X une variable aléatoire discrète de support $X(\Omega)=\{x_0, x_1, x_2, ...\}$ et *f* une fonction continue. Si la série $\sum_k f(x_k)P(X=x_k)$ converge absolument, la variable aléatoire $f(X)$ admet une espérance et

$$E(f(X))=\sum_k f(x_k)P(X=x_k) \quad .$$

Exemple type : $E(X^2)=\sum_k k^2 P(X=k)$ lorsque cette série converge.

<u>Définition :</u> Soit X une variable aléatoire discrète admettant une espérance.

On dit que X *admet une variance* si la variable aléatoire $(X-E(X))^2$ a une espérance.

Dans ce cas, *la variance de* X est définie par $V(X)=E\left[(X-E(X))^2\right]$

et *l'écart type de* X est défini par $\sigma(X)=\sqrt{V(X)}$.

Interprétation :

La variance et l'écart type de X, mesurent les « *écarts de X à sa moyenne* » , ou encore la *dispersion* de X.

<u>Théorème de Koenig-Huygens :</u> Une variable aléatoire discrète X admet une variance si et seulement si X^2 admet une espérance, et dans ce cas :

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

Exemple : Pour X = maximum de 3 dés à 4 faces.

$$E(X^2)= 1^2\times\frac{1}{64}+2^2\times\frac{7}{64}+3^2\times\frac{19}{64}+4^2\times\frac{37}{64}=\frac{792}{64}=\frac{99}{8}$$

Donc $V(X)= E(X^2)-E(X)^2=\frac{99}{8}-\left(\frac{55}{16}\right)^2=\frac{143}{256}$ et $\sigma(X)=\sqrt{\frac{143}{256}}$

Exercice : On considère la variable aléatoire X définie sur $\mathbb{N}^*$ par $P(X=k)=\left(\frac{1}{2}\right)^k$.

1. Justifier que l'on a bien défini une loi de probabilités.

Manifestement $\left(\frac{1}{2}\right)^k\geq 0$ **. Calculons** $\sum_{k=1}^{+\infty}\left(\frac{1}{2}\right)^k=\sum_{k=0}^{+\infty}\left(\frac{1}{2}\right)^k-1=\frac{1}{1-\frac{1}{2}}-1$ **car -1 < ½ < 1**

Donc $\sum_{k=1}^{+\infty}\left(\frac{1}{2}\right)^k=\frac{1}{\frac{1}{2}}-1=2-1=1$ **. on a bien défini une loi de probabilités.**

2. Justifier que X admet une espérance et calculer E(X).

On étudie la série $\sum_{k=1}^{+\infty} |k\,P(X=k)| = \sum_{k=1}^{+\infty} k\left(\frac{1}{2}\right)^k = \frac{1}{2}\sum_{k=1}^{+\infty} k\left(\frac{1}{2}\right)^{k-1} = \frac{1}{2}\times\frac{1}{\left(1-\frac{1}{2}\right)^2}$ **car -1 < ½ <1**

Ainsi $\sum_{k=1}^{+\infty} |k\,P(X=k)| = \frac{1}{2}\times\frac{1}{\left(\frac{1}{2}\right)^2} = \frac{1}{2}\times 4 = 2$.

On conclut que X admet une espérance et que $E(X) = \sum_{k=1}^{+\infty} k\,P(X=k) = 2$

3. Justifier que X admet une variance et calculer V(X).

On étudie la série $\sum_{k=1}^{+\infty} k^2 P(X=k) = \sum_{k=1}^{+\infty} k^2\left(\frac{1}{2}\right)^k = \sum_{k=1}^{+\infty} [k(k-1)+k]\left(\frac{1}{2}\right)^k$ **en notant que** $k^2 = k(k-1) + k$**.**

Ainsi $\sum_{k=1}^{+\infty} k^2 P(X=k) = \sum_{k=1}^{+\infty} k(k-1)\left(\frac{1}{2}\right)^k + \sum_{k=1}^{+\infty} k\left(\frac{1}{2}\right)^k = \left(\frac{1}{2}\right)^2 \sum_{k=1}^{+\infty} k(k-1)\left(\frac{1}{2}\right)^{k-2} + \frac{1}{2}\sum_{k=1}^{+\infty} k\left(\frac{1}{2}\right)^{k-1}$

Comme 1 < ½ <1 on peut appliquer les formules des séries géométriques :

$$\sum_{k=1}^{+\infty} k^2 P(X=k) = \left(\frac{1}{2}\right)^2 \times \frac{2}{\left(1-\frac{1}{2}\right)^3} + \frac{1}{2}\frac{1}{\left(1-\frac{1}{2}\right)^2} = \frac{1}{4}\times 2\times 8 + \frac{1}{2}\times 4 = 6$$

Ainsi $E(X^2) = 6$

Par la formule de Konig-Huijgens, X admet une variance et $V(X) = E(X^2) - E(X)^2 = 6 - 2^2 = 2$

<u>Propriétés :</u> Soit X une variable aléatoire discrète, et $(a,b)\in\mathbb{R}^2$.

- Si X admet une espérance, alors $aX+b$ admet une espérance et $E(aX+b) = a\,E(X)+b$.
- Si X admet une variance, alors $aX+b$ admet une variance et $V(aX+b) = a^2 V(X)$.

3. Lois discrètes usuelles.

Loi uniforme $U([[a,b]])$

Scénario type : **Choix uniforme d'un entier entre deux valeurs a et b.**

Support : **X(Ω) = [[a , b]]**

Loi : $P'(X=k)=\dfrac{1}{b-a+1}$ pour $k\in[[a,b]]$

Histogramme :

Espérance :

Variance :

Loi binomiale $B(n,p)$

Scénario type :

Support :

Loi :

Histogramme :

Espérance :

Variance :

Loi géométrique $G(p)$

Scénario type :

Support : Loi :

Histogramme :

Espérance : Variance :

Loi de Poisson $P(\lambda)$

Scénario type :

Support : Loi :

Histogramme :

Espérance : Variance :

Théorème : Approximation de la loi binomiale par une loi de Poisson

Si X suit une loi Binomiale $B(n, p)$ avec $n \geq 50$, $p \leq 0{,}1$ et $n p \leq 10$,

alors X suit approximativement une loi de Poisson $P(\lambda)$ de paramètre $\lambda = n p$.

Remarque : Nous reviendrons sur ce théorème dans un chapitre ultérieur, en définissant rigoureusement la notion d'*approximation* (et de *convergence*) de variables aléatoires.