Variables aléatoires réelles discrètes

Dans tout ce chapitre, on se place dans un espace probabilisé $(\Omega, \mathcal{F}, P)$.
Les variables discrètes ont été définies dans le chapitre précédent.

1 Loi d'une variable discrète

1.1 Généralités sur les variables discrètes

Exemple 1 :
On effectue une infinité de lancers d'un dé équilibré. On introduit la variable aléatoire X qui donne le nombre de lancers nécessaires à l'obtention du premier 1, si l'on obtient au moins un 1, et 0 si l'on n'obtient jamais 1. Déterminer $X(\Omega)$.

Remarque Autant il n'est pas toujours possible de définir Ω, autant il est indispensable de préciser $X(\Omega)$ dès qu'une variable aléatoire X est introduite.

`Théorème`
Soit X une variable aléatoire discrète sur $(\Omega, \mathcal{F})$. Alors la famille $([X = k])_{k \in X(\Omega)}$ est un <u>système complet d'événements</u>, appelé système complet d'événements associé à la variable X.

$\longrightarrow$ en effet,

`Proposition Opérations sur les variables aléatoires`
Soit X, Y deux variables aléatoires discrètes sur $(\Omega, \mathcal{F})$ et soit $\lambda \in \mathbb{R}$.
Alors $X + Y$, λX, XY, $\max(X, Y)$, $\min(X, Y)$ sont des variables aléatoires discrètes sur $(\Omega, \mathcal{F})$.
Plus généralement, pour toute fonction f définie sur $X(\Omega)$, $f(X)$ est une variable aléatoire discrète sur $(\Omega, \mathcal{F})$.

$\longrightarrow$ $Z = min(X, Y)$ est la variable aléatoire définie pour tout $\omega \in \Omega$ par $Z(\omega) = min(X(\omega), Y(\omega))$.
Plus généralement, si $Z = f(X)$, alors pour tout $\omega \in \Omega$, $Z(\omega) = f(X(\omega))$.

1.2 Loi de probabilité

Définition
Soit X une variable discrète sur $(\Omega, \mathcal{F}, P)$.
On appelle loi de X (ou loi de probabilité de X) l'application $\boxed{\begin{array}{rl} f_X : X(\Omega) & \to \mathbb{R} \\ x & \mapsto P(X = x) \end{array}}$

$\longrightarrow$ $\boxed{\text{en pratique, déterminer la loi de } X \text{ c'est préciser } X(\Omega)\text{, puis donner } P(X = k) \text{ pour tout } k \in X(\Omega)}$.

Remarque
La représentation graphique d'une loi d'une variable aléatoire est un $\boxed{\text{diagramme en bâtons}}$: en abscisse, les valeurs prises et en ordonnée, les probabilités d'apparition.

Retour exemple 1 : Loi de X ?

Propriétés importantes d'une loi de probabilité :
Soit X une variable aléatoire discrète. Alors

1. $\forall k \in X(\Omega)$, $P(X = k) \geq 0$
2. La somme $\sum\limits_{k \in X(\Omega)} P(X = k)$ vaut 1 (sous-entendu la série converge, dans le cas où $X(\Omega)$ est infini).

Réciproquement :
`Théorème`
Soit I un ensemble fini ou dénombrable, $(x_k)_{k \in I}$ une famille de réels distincts et $(p_k)_{k \in I}$ une famille de réels positifs vérifiant $\sum\limits_{k \in I} p_k = 1$ (sous-entendu la série converge, si I est infini).
Alors il existe une variable aléatoire, définie sur un espace probabilisé, vérifiant :
$$X(\Omega) = \{x_k, k \in I\} \text{ et pour tout } k \in I,\ P(X = x_k) = p_k.$$
$\longrightarrow$ Autrement dit, les x_k sont les valeurs prises par X, et les p_k les probabilités d'apparition.

Exemple :
Soit X une variable aléatoire définie sur un certain espace probabilisé $(\Omega, \mathcal{F})$ telle que $X(\Omega) = \{2k + 1, k \in \mathbb{N}\}$

et pour tout $k \in \mathbb{N}$, $P(X = 2k+1) = \frac{a}{3^{k+1}}$.
Déterminer a afin que l'on définisse bien une loi de probabilité.

1.3 Indépendance de variables discrètes

`Proposition`

Soit X et Y deux variables aléatoires <u>discrètes</u> définies sur un même espace $(\Omega, \mathcal{F}, P)$.
Alors X et Y sont indépendantes si et seulement si

pour tout $(i,j) \in X(\Omega) \times Y(\Omega)$, $P((X = i) \cap (Y = j)) = P(X = i) \times P(Y = j)$.

Plus généralement :

— Soit une famille finie $(X_1, X_2, ... ,X_n)$ de variables aléatoires définies sur un même espace $(\Omega, \mathcal{F}, P)$. Ces variables sont <u>mutuellement indépendantes</u> ssi pour tout $(x_1, x_2, ..., x_n) \in X_1(\Omega) \times X_2(\Omega) \times ... \times X_n(\Omega)$,
$$P((X_1 = x_1) \cap (X_2 = x_2) \cap ... \cap (X_n = x_n)) = \prod_{k=1}^{n} P(X_k = x_k).$$

— Enfin, un rappel : soit $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires discrètes définies sur un même espace $(\Omega, \mathcal{F}, P)$. $(X_n)_{n\in\mathbb{N}}$ est une suite de variables indépendantes si pour tout ensemble d'indices $I \subset \mathbb{N}$ <u>fini</u>, les variables $(X_k)_{k\in I}$ sont mutuellement indépendantes.

Propriétés de l'indépendance mutuelle :

— Toute sous-famille d'une famille de variables indépendantes est indépendante.

— <u>Lemme des coalitions</u> : soit une famille $(X_1, ..., X_n, X_{n+1}, ..., X_{n+p})$ de variables aléatoires mutuellement indépendantes. Alors pour toutes fonctions u et v, définies sur $\mathbb{R}^n$ et $\mathbb{R}^p$, les variables $u(X_1, ..., X_n)$ et $v(X_{n+1}, ...X_{n+p})$ sont indépendantes.

— Le lemme des coalitions peut se généraliser à une partition plus grande des variables : en particulier si $(X_1,...X_n)$ est une famille de variables mutuellement indépendantes, pour toutes fonctions u_1, ..., u_n défnies sur $\mathbb{R}$, les variables $u_1(X_1)$, ..., $u_n(X_n)$ sont encore mutuellement indépendantes.

2 Moments d'une variable aléatoire discrète

2.1 Espérance

Définition Soit X une variable aléatoire discrète sur $(\Omega, \mathcal{F}, P)$.
On dit que X admet une espérance si $X(\Omega)$ est fini ou si la série $\sum_{k\in X(\Omega)} kP(X = k)$ converge absolument. On appelle alors espérance de X le réel $E(X) = \sum_{k\in X(\Omega)} kP(X = k)$ (qui est la somme de la série, lorsque $X(\Omega)$ est infini).

Retour Exemple 1 : Montrer que X admet une espérance et la calculer.

Propriétés de l'espérance

`Proposition`

Soit X et Y deux variables discrètes sur $(\Omega, \mathcal{F}, P)$ $\boxed{\text{admettant une espérance}}$ et a, b deux nombres réels. Alors

1. `Linéarité de l'espérance`
   (a) la variable $aX + b$ <u>admet une espérance</u> et $E(aX + b) = aE(X) + b$.
   (b) la variable $X + Y$ <u>admet une espérance</u> et $E(X + Y) = E(X) + E(Y)$.
2. `Positivité` : Si $X \geq 0$ (signifie X ne prend que des valeurs positives), alors $E(X) \geq 0$.
3. `Croissance` : Si $X \leq Y$, alors $E(X) \leq E(Y)$.
4. Si de plus, X et Y sont indépendantes, alors XY <u>admet une espérance</u> et $E(XY) = E(X)E(Y)$.

— `Preuve du 2.` On sait par hypothèse que X admet une espérance, donc $E(X) =$

— `Preuve du 3.`

`Proposition` `Théorème de transfert admis`

$\boxed{\text{Soit } X \text{ une variable discrète sur } (\Omega,\mathcal{A},P) \text{, et } f \text{ une fonction réelle définie sur } X(\Omega)\text{.}}$

Soit X une variable discrète sur $(\Omega, \mathcal{A}, P)$, et f une fonction réelle définie sur $X(\Omega)$. Alors la variable $f(X)$ admet une espérance si et seulement si $X(\Omega)$ est fini ou si la série $\sum\limits_{k\in X(\Omega)} f(k)P(X=k)$ converge absolument, et dans ce cas, $E(f(X)) = \sum\limits_{k\in X(\Omega)} f(k)P(X=k)$ (qui est la somme de la série, lorsque $X(\Omega)$ est infini).

Cas particulier : $f(x) = x^2$.

X^2 admet une espérance si $X(\Omega)$ est fini ou si la série $\sum\limits_{k\in X(\Omega)} k^2 P(X=k)$ converge. Le cas échéant, $E(X^2) = \sum\limits_{k\in X(\Omega)} k^2 P(X=k)$. *(ici, inutile de justifier la convergence absolue car tous les termes sont positifs)*

2.2 Variance

Définition

Soit X une var discrète sur $(\Omega, \mathcal{F}, P)$. Sous-réserve d'existence, on appelle

1. moment d'ordre 2 de X le réel $E(X^2)$.
2. variance de X le réel $V(X) = E[(X - E(X))^2]$ (mesure l'écart entre X et sa moyenne)
3. écart-type de X le réel $\sigma(X) = \sqrt{V(X)}$

Remarque

X peut admettre une espérance sans admettre de moment d'ordre 2. Mais si X admet un moment d'ordre 2, alors X admet une espérance, et donc par contraposée, si X n'admet pas d'espérance alors X n'admet pas non plus de moment d'ordre 2.

$\longrightarrow$ `idée de la preuve`, dans le cas où $X(\Omega) = \mathbb{N}$

`Proposition Formule de Koenig-Huygens`

X admet un moment d'ordre 2 ssi X admet une variance et dans ce cas, $V(X) = E(X^2) - (E(X))^2$.

$\longrightarrow$ `preuve` dans le cas $X(\Omega) = \mathbb{N}$, en notant $m = E(X)$.

Propriétés de la variance

`Proposition`

Soit X et Y deux variables discrètes sur $(\Omega, \mathcal{F}, P)$ $\boxed{\text{admettant une variance}}$ et a, b deux nombres réels. Alors

1. $V(X) \geq 0$.
2. $V(X) = 0 \Leftrightarrow X$ est une variable aléatoire constante sur Ω presque sûrement.
3. La variable $aX + b$ <u>admet une variance</u> et $\boxed{V(aX+b) = a^2V(X)}$.
4. $\boxed{\text{Si de plus } X \text{ et } Y \text{ sont indépendantes, alors } V(X+Y) = V(X) + V(Y)}$.

$\longrightarrow$ `preuve 1. et 3.`

Définition

Une variable X est dite `centrée` lorsque $E(X) = 0$ et `réduite` lorsque $V(X) = 1$.

`Proposition`

Soit X une variable admettant une variance non-nulle (donc X n'est pas constante). Alors,
la variable $X^* = \dfrac{X - E(X)}{\sigma(X)}$ est une variable centrée réduite, appelée variable centrée réduite associée à X.

$\longrightarrow$ En effet,

3 Lois usuelles discrètes

3.1 Rappel des lois usuelles discrètes finies

Loi	Notation	Valeurs prises	Probabilités d'apparitions	Espérance	Variance
Loi certaine		$X(\Omega)=\{a\}$	$P(X=a)=1$	$E(X)=$	$V(X)=$
Loi de Bernoulli	$X\hookrightarrow\mathcal{B}(p)$	$X(\Omega)=\{0,1\}$	$P(X=1)=p$, $P(X=0)=1-p$	$E(X)=$	$V(X)=$
Loi Binomiale	$X\hookrightarrow\mathcal{B}(n,p)$	$X(\Omega)=[\![0,n]\!]$	$\forall k\in X(\Omega)$, $P(X=k)=\binom{n}{k}p^k(1-p)^{n-k}$	$E(X)=$	$V(X)=$
Loi uniforme	$X\hookrightarrow\mathcal{U}([\![1,n]\!])$	$X(\Omega)=[\![1,n]\!]$	$\forall k\in X(\Omega)$, $P(X=k)=\frac{1}{n}$	$E(X)=$	

Remarque : La loi uniforme discrète se généralise sur tout intervalle d'entiers : $X\hookrightarrow\mathcal{U}([\![a,b]\!]$ si $X(\Omega)=[\![a,b]\!]$ et pour tout $k\in X(\Omega)$, $P(X=k)=\frac{1}{b-a+1}$.

Citer un cas typique de la loi de Bernoulli, de la loi uniforme et de la loi binomiale :

$\longrightarrow$ `Preuve des probabilités d'apparition de la loi binomiale`

Remarque
La loi binomiale peut être vue comme une somme de n bernoullis. Autrement dit, lors d'expériences identiques et indépendantes de même probabilité de succès $P(S)$, $\sum\limits_{k=1}^{n}\mathbb{1}_{S_i}$ suit la loi binomiale de paramètre n et $P(S)$.

3.2 Lois usuelles discrètes infinies

3.2.1 Loi géométrique

Définition
On effectue une infinité d'épreuves mutuellement indépendantes. Chaque épreuve a deux issues : le succès de probabilité $p\in]0,1[$ et l'échec de probabilité $(1-p)$, et on introduit la variable X égale au rang d'apparition du premier succès, càd au nombre d'épreuves nécessaires à l'obtention du premier succès.
On dit alors que X suit la loi géométrique de paramètre p et on note : $X\hookrightarrow\mathcal{G}(p)$.
$\longrightarrow$ *X est le temps d'attente du premier succès dans un processus de Bernoulli sans mémoire*

`Théorème` Soit $p\in]0,1[$. Si une variable aléatoire X suit la loi géométrique de paramètre p alors
$$\boxed{X(\Omega)=\mathbb{N}^*,\text{ et }\forall n\in\mathbb{N}^*,\ P(X=n)=p(1-p)^{n-1}}$$

$\longrightarrow$ `preuve`

Remarque Soit le modèle suivant :
On effectue une succession d'épreuves mutuellement indépendantes, jusqu'à l'obtention du premier succès.
Chaque épreuve a deux issues : le succès de probabilité $p\in]0,1[$ et l'échec de probabilité $(1-p)$.
X est la variable aléatoire égale au rang d'apparition du premier succès, c'est-à-dire au nombre d'épreuves effectuées. On pourrait démontrer de la même manière (en utilisant la formule des probabilités composées au lieu de la définition de la mutuelle indépendance) que dans ce cas, on a aussi $X\hookrightarrow\mathcal{G}(p)$.

Espérance et Variance :
Si $X\hookrightarrow\mathcal{G}(p)$ alors X admet une espérance et une variance et $\boxed{E(X)=\frac{1}{p},\ V(X)=\frac{1-p}{p^2}}$.

$\longrightarrow$ `preuve`

Propriété d'absence de mémoire :
Soit $X\hookrightarrow\mathcal{G}(p)$.
Montrer que $\boxed{\text{pour tout }(n,k)\in\mathbb{N}^2,\ P_{(X>n)}(X>n+k)=P(X>k)}$, puis interpréter.

3.2.2 Loi de Poisson

Soit $\lambda > 0$. X suit la loi de Poisson de paramètre λ, notée $\boxed{\mathcal{P}(\lambda)}$ si $\boxed{X(\Omega) = \mathbb{N} \text{ et } \forall n \in \mathbb{N},\ P(X = n) = \dfrac{\lambda^n}{n!}e^{-\lambda}}$

Remarque : on définit bien une loi de probabilité car pour tout $n \in \mathbb{N}$, $\frac{\lambda^n}{n!}e^{-\lambda} \geq 0$, et $\sum\limits_{n=0}^{+\infty} \dfrac{\lambda^n}{n!}e^{-\lambda} = e^{-\lambda}e^{\lambda} = 1$.

<u>*Espérance et Variance*</u>

Si $X \hookrightarrow \mathcal{P}(\lambda)$ alors X admet une espérance et une variance et $\boxed{E(X) = V(X) = \lambda}$.

$\longrightarrow$ `preuve`

4 Fonction de répartition d'une variable discrète

Les généralités ont été revues dans le chapitre précédent (définition, premières propriétés). Dans le cas des variables discrètes, la fonction de répartition a une allure particulière qui généralise l'allure étudiée dans le cas des variables aléatoires finies en première année.

Exemples : déterminer la fonction de répartition d'une variable certaine égale à a, d'une variable de loi de Bernoulli $\mathcal{B}(p)$, et d'une variable de loi $\mathcal{B}(2, \frac{1}{4})$.

`Proposition Rappel`

Soit X une variable <u>finie</u> sur $(\Omega, \mathcal{P}(\Omega), P)$ avec $X(\Omega) = \{x_1, ..., x_n\}$ et $x_1 < x_2 < ... < x_n$.
Sa fonction de répartition F_X est constante par morceaux (en escalier) et ses points de discontinuités sont les $x_i, i \in [\![1, n]\!]$ (les valeurs prises par X). Plus précisément, $\forall t \in \mathbb{R}$,

$$F_X(t) = \begin{cases} 0 & \text{si } t < x_1 \\ P(X = x_1) & \text{si } x_1 \leq t < x_2 \\ P(X = x_1) + P(X = x_2) & \text{si } x_2 \leq t < x_3 \\ P(X = x_1) + P(X = x_2) + P(X = x_3) & \text{si } x_3 \leq t < x_4 \\ .... & \\ P(X = x_1) + P(X = x_2) + \cdot + P(X = x_{n-1}) & \text{si } x_{n-1} \leq t < x_n \\ 1 & \text{si } x_n \leq t \end{cases}$$

`Théorème`

Soit X une variable discrète infinie telle que $X(\Omega) = \{x_0, x_1, ..., x_n, ...\}$ où $(x_n)_{n \in \mathbb{N}}$ strictement croissante. Alors la fonction de répartition de X, F_X, est une fonction constante sur chaque intervalle $[x_k, x_{k+1}[$.

$\longrightarrow$ F_X peut être visualisée comme une fonction en escalier "généralisée" (avec une infinité de palliers).

Pourquoi l'hypothèse $(x_n) \nearrow$? Soit X de loi : $X(\Omega) = \{\frac{1}{n}, n \in \mathbb{N}^*\}$ et $\forall n \geq 1$, $P(X = \frac{1}{n}) = \frac{1}{2^n}$. La fonction de répartition est bien sûr constante sur chaque intervalle de type $[\frac{1}{n+1}, \frac{1}{n}[$, mais les palliers se "tassent " en 0.

> A garder en mémoire :
> Si X est une variable discrète, la fonction de répartition de X est constante par morceaux (généralisée), avec des sauts en les valeurs prises par X. Donc déterminer F_X c'est déterminer $F_X(k)$ pour tout $k \in X(\Omega)$.

En pratique :

1. La connaissance de la loi de X permet la détermination de F_X.
   En effet, dans le cas où $X(\Omega) \subset \mathbb{N}$, pour tout $n \in \mathbb{N}$, $(X \leq n) = \bigcup\limits_{k=0}^{n} (X = k)$, réunion d'événements 2 à 2 incompatibles d'où $F_X(n) = P(X \leq n) = \sum\limits_{k=0}^{n} P(X = k)$.
2. Réciproquement : la connaissance de F_X permet la détermination de la loi de X.
   En effet dans le cas où $X(\Omega) \subset \mathbb{N}$, pour tout $n \in \mathbb{N}$, $(X \leq n) = (X = n) \cup (X \leq n-1)$ $(*)$ réunion de 2 événements incompatibles, d'où $P(X = n) = P(X \leq n) - P(X \leq n-1) = F_X(n) - F_X(n-1)$.

Remarque Suivre l'énoncé ! Si on vous demande de déterminer pour tout $k \in X(\Omega) \subset \mathbb{N}$, $P(X > k)$, puis d'en déduire la loi de X, adapter la relation $(*)$: pour tout $k \in X(\Omega)$, $P(X = k) =$

De même, si les probabilités connues sont les $P(X \geq k)$, alors $P(X = k) =$