Chapitre 3
Variables aléatoires discrètes

1 Variables aléatoires discrètes à une dimension

Soient $f : E \to F$ et $A \subset F$. Le sous-ensemble $f^{-1}(A) \subset E$ est défini par

$$f^{-1}(A) = \{x \in E \mid f(x) \in A\}.$$

Définition 1. On appelle variable aléatoire discrète (sur $(\Omega, \mathcal{T}, \mathbb{P})$) une application $X : \Omega \to \mathbb{N}$ (ou $\mathbb{Z}$) telle que pour tout $n \in \mathbb{N}$ (ou $\mathbb{Z}$), $X^{-1}(\{n\}) \in \mathcal{T}$.

On notera $\{X = n\} = X^{-1}(\{n\})$ pour tout $n \in X(\Omega)$.
On pourra parfois considérer un autre ensemble discret D au lieu de $\mathbb{N}$, c'est-à-dire un ensemble D en bijection avec un sous-ensemble de $\mathbb{N}$.

Loi d'une variable discrète

Définition 2. On appelle loi de la variable aléatoire X l'ensemble des données suivantes :

1. $X(\Omega)$;
2. Pour tout $n \in X(\Omega)$, $\mathbb{P}(X = n) = \mathbb{P}(X^{-1}(\{n\}))$.

On notera généralement

$$\mathbb{P}(X \in A) = \mathbb{P}(X^{-1}(A))$$

pour tout $A \in \mathcal{P}(X(\Omega))$.

Théorème 1. *Pour tout $A \in \mathcal{P}(X(\Omega))$, $\mathbb{P}(X \in A) = \sum_{n \in A} \mathbb{P}(X = n)$.*

Exemples de lois discrètes

Exemple 1. Loi de Bernoulli. On dit qu'une variable aléatoire X suit une loi de Bernoulli de paramètre $p \in]0, 1[$ (et on note $X \sim \mathcal{B}er(p)$) si

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

Exemple 2. Loi binomiale. On dit qu'une variable aléatoire X suit une loi binomiale de paramètres $(n, p) \in \mathbb{N}^* \times]0, 1[$ (et on note $X \sim \mathcal{B}in(n, p)$) si

1. $X(\Omega) = \{0, 1, ..., n\}$ et
2. pour tout $k \in \{0, 1, ..., n\}$, $\mathbb{P}(X = k) = C_n^k p^k (1-p)^{n-k}$.

Elle représente le nombre de succès lorsque l'on fait n fois de façon indépendante la même expérience aléatoire dont la probabilité de succès est p.

Exemple 3. Loi géométrique. On dit qu'une variable aléatoire X suit une loi géométrique de paramètre $p \in]0, 1[$ (et on note $X \sim \mathcal{G}(p)$) si

1. $X(\Omega) = \mathbb{N}^*$ et
2. pour tout $k \in \mathbb{N}^*$, $\mathbb{P}(X = k) = (1-p)^{k-1} p$.

Elle représente la position du premier succès lorsque l'on répète (une infinité de fois ou jusqu'au premier succès) de façon indépendante la même expérience aléatoire dont la probabilité de succès est p.

Exemple 4. Loi de Poisson. On dit qu'une variable aléatoire X suit une loi de Poisson de paramètre $\lambda > 0$ (et on note $X \sim \mathcal{P}oisson(\lambda)$) si

1. $X(\Omega) = \mathbb{N}$ et
2. pour tout $k \in \mathbb{N}$, $\mathbb{P}(X = k) = e^{-\lambda} \dfrac{\lambda^k}{k!}$.

Espérance

Définition 3. On dit que X est intégrable si

$$\sum_{k \in X(\Omega)} |k| \mathbb{P}(X = k) < \infty.$$

Dans ce cas, on définit l'espérance de X par

$$\mathbb{E}(X) = \sum_{k \in X(\Omega)} k \mathbb{P}(X = k).$$

Exemples :

Si $X \sim \mathcal{B}er(p)$, alors

$$\begin{aligned}
\mathbb{E}(X) &= \sum_{k=0}^{1} k \mathbb{P}(X = k) = 0 \times \mathbb{P}(X = 0) + 1 \times \mathbb{P}(X = 1) \\
&= 0 \times (1-p) + 1 \times p = p.
\end{aligned}$$

Si $X \sim \mathcal{P}oisson(\lambda)$, alors

$$\begin{aligned}
\mathbb{E}(X) &= \sum_{k=0}^{\infty} k \mathbb{P}(X = k) = \sum_{k=0}^{\infty} k e^{-\lambda} \frac{\lambda^k}{k!} = \sum_{k=1}^{\infty} k e^{-\lambda} \frac{\lambda^k}{k!} \\
&= \sum_{k=1}^{\infty} e^{-\lambda} \frac{\lambda^k}{(k-1)!} \underset{k=j+1}{\overset{j=k-1}{=}} \sum_{j=0}^{\infty} e^{-\lambda} \frac{\lambda^{j+1}}{j!} = \lambda e^{-\lambda} \sum_{j=0}^{\infty} \frac{\lambda^j}{j!} = \lambda.
\end{aligned}$$

Variance et écart-type

Définition 4. Si X^2 est intégrable, on définit la variance X par

$$\mathrm{Var}(X) = \mathbb{E}((X - \mathbb{E}(X))^2) = \mathbb{E}(X^2) - \mathbb{E}(X)^2$$

et son écart-type par $\sigma(X) = \sqrt{\mathrm{Var}(X)}$.

Théorème 2. *La variance vérifie les propriétés suivantes.*

1. $Var(X) \geq 0$.
2. $Var(X) = 0$ *si et seulement si* X *est constante.*
3. *Pour tout* $a, b \in \mathbb{R}$, $Var(aX + b) = a^2 Var(X)$.

Calcul de la variance :

Théorème 3 (Théorème de transfert)**.** *Soit* g *une fonction sur* $\mathbb{N}$. *La variable aléatoire* $g(X) = g \circ X$ *est intégrable si et seulement si* $\sum_{n \in X(\Omega)} |g(n)|\mathbb{P}(X = n) < \infty$. *De plus, si* $g(X)$ *est intégrable, alors*

$$\mathbb{E}(g(X)) = \sum_{n \in X(\Omega)} g(n)\mathbb{P}(X = n).$$

1. On utilise $\mathrm{Var}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2$.
2. L'espérance au carré $\mathbb{E}(X)^2$ se calcule facilement.
3. La quantité $\mathbb{E}(X^2)$ se calcule avec le théorème de transfert :

$$\mathbb{E}(X^2) = \sum_{n \in X(\Omega)} n^2 \mathbb{P}(X = n)$$

   avec la fonction $g(x) = x^2$.

Exemple :

Par exemple, si $X \sim Ber(p)$, on a $\mathbb{E}(X)^2 = p^2$ (car on a vu que $\mathbb{E}(X) = p$), et

$$\mathbb{E}(X^2) = 1^2\mathbb{P}(X = 1) + 0^2\mathbb{P}(X = 0) = p.$$

Ainsi : $Var(X) = p - p^2 = p(1 - p)$.

2 Vecteurs aléatoires discrets

Définition 5. Soient $d \in \mathbb{N}^*$. Un vecteur aléatoire discret $(X_1, X_2, ..., X_d)$ est un vecteur dont les composantes X_1, X_2, ..., X_d sont d variables aléatoires discrètes. La loi du vecteur aléatoire discret $(X_1, X_2, ..., X_d)$ est définie par les données suivantes.

1. $X_1(\Omega)$, $X_2(\Omega)$, ..., $X_d(\Omega)$.
2. Pour tout $(n_1, ..., n_d) \in X_1(\Omega) \times X_2(\Omega) \times ... \times X_d(\Omega)$,

$$\begin{aligned}&\mathbb{P}(X_1 = n_1, X_2 = n_2, ..., X_d = n_d)\\ &\qquad = \mathbb{P}(\{X_1 = n_1\} \cap \{X_2 = n_2\} \cap ... \cap \{X_d = n_d\}).\end{aligned}$$

Lois marginales

Définition 6. On appelle lois marginales d'un vecteur aléatoire discret $(X_1, X_2, ..., X_d)$ les lois de X_1, X_2, ... et X_d. Pour un couple aléatoire discret (X, Y), les lois marginales sont données par $X(\Omega)$, $Y(\Omega)$, et les formules suivantes :

$$\mathbb{P}(X = n) = \sum_{m \in Y(\Omega)} \mathbb{P}(X = n, Y = m) \qquad \text{pour tout } n \in X(\Omega),$$
$$\mathbb{P}(Y = m) = \sum_{n \in X(\Omega)} \mathbb{P}(X = n, Y = m) \qquad \text{pour tout } m \in Y(\Omega).$$

Lois conditionnelles

Définition 7. Pour un couple aléatoire discret (X, Y) et $m \in Y(\Omega)$ tel que $\mathbb{P}(Y = m) > 0$, on appelle loi conditionnelle de X sachant $Y = m$ l'ensemble des données suivantes.

1. $X(\Omega)$.
2. Pour tout $n \in X(\Omega)$,

$$\mathbb{P}(X = n \,|\, Y = m) = \frac{\mathbb{P}(X = n, Y = m)}{\mathbb{P}(Y = m)}.$$

On peut aussi écrire

$$\mathbb{P}(X = n, Y = m) = \mathbb{P}(X = n \,|\, Y = m)\mathbb{P}(Y = m).$$

Indépendance

Définition 8. On dit que deux variables aléatoires discrètes X et Y sont indépendantes si pour tout $(n, m) \in X(\Omega) \times Y(\Omega)$,

$$\mathbb{P}(X = n, Y = m) = \mathbb{P}(X = n)\mathbb{P}(Y = m).$$

On dit que d variables aléatoires discrètes X_1, X_2, ... et X_d sont mutuellement indépendantes si pour tout $(n_1, ..., n_d) \in X_1(\Omega) \times ... \times X_d(\Omega)$,

$$\mathbb{P}(X_1 = n_1, X_2 = n_2, ..., X_d = n_d) = \mathbb{P}(X_1 = n_1) \times ... \times \mathbb{P}(X_d = n_d).$$

On dit que d variables aléatoires discrètes X_1, X_2, ... et X_d sont deux à deux indépendantes si pour tout $(j, k) \in \{1, ..., d\}^2$ tel que $j \neq k$ et tout $(n, m) \in X_j(\Omega) \times X_k(\Omega)$,

$$\mathbb{P}(X_j = n, X_k = m) = \mathbb{P}(X_j = n)\mathbb{P}(X_k = m).$$

Exercice :

Exercice 1. Soient X et Y deux variables aléatoires indépendantes de lois de Poisson de paramètres λ et μ respectivement. Quelle est la loi de $X+Y$?

On rappelle que $X(\Omega) = Y(\Omega) = \mathbb{N}$ et que pour tout $n \in \mathbb{N}$,

$$\mathbb{P}(X = n) = e^{-\lambda}\frac{\lambda^n}{n!} \quad \text{et} \quad \mathbb{P}(Y = n) = e^{-\mu}\frac{\mu^n}{n!}.$$

On a $(X+Y)(\Omega) = \mathbb{N}$.

De plus, pour tout $n \in \mathbb{N}$,

$$\begin{aligned}
\mathbb{P}(X+Y=n) &= \sum_{k=0}^{n} \mathbb{P}(X=k, Y=n-k) \\
&= \sum_{k=0}^{n} \mathbb{P}(X=k)\mathbb{P}(Y=n-k) \text{ (indépendance)} \\
&= e^{-\lambda-\mu} \sum_{k=0}^{n} \frac{\lambda^k}{k!}\frac{\mu^{n-k}}{(n-k)!} = e^{-\lambda-\mu}\frac{1}{n!}\sum_{k=0}^{n}\frac{\lambda^k n!}{k!}\frac{\mu^{n-k}}{(n-k)!} \\
&= e^{-\lambda-\mu}\frac{1}{n!}\sum_{k=0}^{n} C_n^k \lambda^k \mu^{n-k} \text{ car } C_n^k = \frac{n!}{k!(n-k)!} \\
&= e^{-\lambda-\mu}\frac{(\lambda+\mu)^n}{n!}
\end{aligned}$$

Donc $X+Y$ suit une loi de Poisson de paramètre $\lambda+\mu$.

Théorème de transfert multidimensionnel

Théorème 4. *Soient g une fonction de d variables et $(X_1, X_2, ..., X_d)$ un vecteur aléatoire discret. Si la variable aléatoire discrète Z définie par $Z = g(X_1, X_2, ..., X_d)$ est intégrable, alors*

$$\begin{aligned}
\mathbb{E}(Z) &= \mathbb{E}(g(X_1, X_2, ..., X_d)) \\
&= \sum_{(n_1,...,n_d)\in X_1(\Omega)\times...\times X_d(\Omega)} g(n_1, ..., n_d)\mathbb{P}(X_1 = n_1, X_2 = n_2, ..., X_d = n_d).
\end{aligned}$$

Caractérisation des variables indépendantes

Théorème 5. *Soit $(X_1, X_2, ..., X_d)$ un vecteur aléatoire discret. Alors $X_1, X_2, ..., X_d$ sont indépendantes si et seulement si pour toutes fonctions bornées $h_1, h_2, ..., h_d$,*

$$\mathbb{E}(h_1(X_1)h_2(X_2)...h_d(X_d)) = \mathbb{E}(h_1(X_1))\mathbb{E}(h_2(X_2))...\mathbb{E}(h_d(X_d)).$$

Preuve : Considérons $d = 2$. Nous ne faisons que la preuve de : si X et Y sont indépendantes, alors $\mathbb{E}(f(X)g(Y)) = \mathbb{E}(f(X))\mathbb{E}(g(Y))$.
On a

$$\begin{aligned}
\mathbb{E}(f(X)g(Y)) &= \sum_{(k,j)\in X(\Omega)\times Y(\Omega)} f(k)g(j)\mathbb{P}(X=k, Y=j) \\
&= \sum_{(k,j)\in X(\Omega)\times Y(\Omega)} f(k)g(j)\mathbb{P}(X=k)\mathbb{P}(Y=j) \\
&= \sum_{k\in X(\Omega)} f(k)\mathbb{P}(X=k) \sum_{j\in Y(\Omega)} g(j)\mathbb{P}(Y=j) \\
&= \mathbb{E}(f(X))\mathbb{E}(g(Y)).
\end{aligned}$$

3 Fonctions génératrices

Définition 9. Soit X une variable aléatoire discrète positive ($X(\Omega) \subset \mathbb{N}$).On définit pour tout $u \in [0, 1]$ la fonction génératrice de X notée G_X par :

$$G_X(u) = \mathbb{E}[u^X] = \sum_{n\in\mathbb{N}} u^n \mathbb{P}(X = n).$$

Théorème 6. *$G_X(u)$ est une série entière en u dont le rayon de convergence est supérieur ou égal à 1. G_X est infiniment dérivable en u, et*

$$G_X^{(n)}(0) = n!\mathbb{P}(X = n).$$

Donc, si X et Y sont deux variables aléatoires discrètes positives, alors X et Y ont même loi si et seulement si $G_X = G_Y$.

Théorème 7. *Soient $X_1, X_2, ..., X_n$ des variables aléatoires discrètes mutuellement indépendantes, alors*

$$G_{X_1+X_2+...+X_n}(u) = \prod_{i=1}^{n} G_{X_i}(u).$$

Preuve : On a

$$\begin{aligned}
G_{X_1+X_2+...+X_n}(u) &= \mathbb{E}(u^{(X_1+X_2+...+X_n)}) \\
&= \mathbb{E}(u^{X_1}u^{X_2}\cdots u^{X_n}) \\
&= \mathbb{E}(u^{X_1})\mathbb{E}(u^{X_2})\cdots\mathbb{E}(u^{X_n}) \text{ (indépendance)} \\
&= G_{X_1}(u)G_{X_2}(u)\cdots G_{X_n}(u).
\end{aligned}$$

Théorème 8. *On a $G'_X(u) = \mathbb{E}(Xu^{X-1})$, et $G'_X(1) = \mathbb{E}(X)$. De même $G_X^{(n)}(1) = \mathbb{E}(X(X-1)(X-2)...(X-(n-1)))$.*

En particulier :

$$\begin{aligned}\mathrm{Var}(X) &= \mathbb{E}(X^2) - \mathbb{E}(X)^2\\ &= \mathbb{E}(X^2) - \mathbb{E}(X) + \mathbb{E}(X) - \mathbb{E}(X)^2\\ &= \mathbb{E}(X(X-1)) + \mathbb{E}(X) - \mathbb{E}(X)^2\\ &= G_X''(1) + G_X'(1) - (G_X'(1))^2.\end{aligned}$$

Exemples

On suppose que $X \sim \mathcal{B}in(n,p)$. On a

$$\begin{aligned}G_X(u) = \mathbb{E}(u^X) &= \sum_{k=0}^{n} u^k \mathbb{P}(X=k) = \sum_{k=0}^{n} u^k C_n^k p^k (1-p)^{n-k}\\ &= (1-p+up)^n\end{aligned}$$

d'après la formule du binôme de Newton $\sum_{k=0}^{n} C_n^k a^k b^{n-k} = (a+b)^n$.

Calcul de $\mathbb{E}(X)$ et $\mathrm{Var}(X)$ avec les fonctions génératrices :
On rappelle les formules suivantes :

$$\mathbb{E}(X) = G_X'(1) \quad \text{et} \quad \mathrm{Var}(X) = G_X''(1) + G_X'(1) - (G_X'(1))^2.$$

On a $G_X'(u) = np(1-p+up)^{n-1}$, donc $G_X'(1) = np$ et $\mathbb{E}(X) = np$.
On a $G_X''(u) = n(n-1)p^2(1-p+up)^{n-2}$, donc $G_X''(1) = n(n-1)p^2$ et $\mathrm{Var}(X) = np(1-p)$.

On suppose que $X \sim \mathcal{P}oisson(\lambda)$. On a

$$G_X(u) = \mathbb{E}(u^X) = \sum_{k=0}^{\infty} u^k \mathbb{P}(X=k) = \sum_{k=0}^{\infty} u^k e^{-\lambda}\frac{\lambda^k}{k!} = e^{\lambda u - \lambda}.$$

Calcul de $\mathbb{E}(X)$ et $\mathrm{Var}(X)$ avec les fonctions génératrices :
On rappelle les formules suivantes :

$$\mathbb{E}(X) = G_X'(1) \quad \text{et} \quad \mathrm{Var}(X) = G_X''(1) + G_X'(1) - (G_X'(1))^2.$$

On a $G_X'(u) = \lambda e^{\lambda u - \lambda}$, donc $G_X'(1) = \lambda$ et $\mathbb{E}(X) = \lambda$.

On a $G_X''(u) = \lambda^2 e^{\lambda u - \lambda}$, donc $G_X''(1) = \lambda^2$ et $\mathrm{Var}(X) = \lambda^2 + \lambda - \lambda^2 = \lambda$.