Chapitre 11

Variable

Hugo Salou MPI*

Dernière mise à jour le 10 janvier 2023

Première partie
Cours

1 Variables aléatoires discrètes

Définition 1:

1. Soit $(\Omega, \mathcal{A})$ un espace probabilisable. Une *variable aléatoire discrète* (*vad*) est une fonction X définie sur l'unitvers Ω telle que
   (a) l'ensemble $X(\Omega)$ des valeurs prises par X est fini ou dénombrable ;
   (b) pour chaque valeur $a \in X(\Omega)$ prise par X, l'ensemble $X^{-1}(\{a\})$ est un événement, noté $(X = a)$. Autrement dit,
   $$\forall a \in X(\Omega), (X = a) = X^{-1}(\{a\}) \in \mathcal{A}.$$
2. Soient $(\Omega, \mathcal{A}, P)$ un espace probabilisé, et X une variable aléatoire discrète. La *loi de probabilité* de X est la fonction
   $$\begin{aligned} X(\Omega) &\longrightarrow [0,1] \\ a &\longmapsto P(X = a). \end{aligned}$$

Rappel:
Soit $f : E \to F$ une fonction. On a
$$\forall A \in \wp(E),\ f(A) = \{y \in F \mid \exists x \in A,\ y = f(x)\},$$
et
$$\forall B \in \wp(F),\ f^{-1}(B) = \{x \in E \mid f(x) \in B\}.$$
Ainsi, $x \in f^{-1}(B) \iff f(x) \in B$.

Exercice 2:
L'univers Ω, ensemble des résultats, est $[\![1,6]\!]^2$. Soit X la variable aléatoire, qui est la somme des deux dés : on a
$$\begin{aligned} X : \Omega &\longrightarrow \mathbb{R} \\ (x,y) &\longmapsto x + y. \end{aligned}$$
Ainsi, $X(\Omega) = [\![2,12]\!]$.

a	2	3	4	5	6	7	8	9	10	11	12
$P(X=a)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

Table 1 – Loi de probabilité de X

Remarque 3:

1. On a $f^{-1}\left(\bigcup_{i\in I} A_i\right) = \bigcup_{i\in I} f^{-1}(A_i)$, et $f^{-1}\left(\bigcap_{i\in I} A_i\right) = \bigcap_{i\in I} f^{-1}(A_i)$.

Exemple 4:
c.f. polycopié

Proposition – Définition 5:
Soient $(\Omega, \mathcal{A}, P)$ un espace probabilisé, et une *vad* $X : \Omega \to E$ à valeurs dans un ensemble E.

1. Pour tout $A \in \wp(E)$, $X^{-1}(A) \in \mathcal{A}$ est un événement noté $(X \in A)$.
2. L'application
   $$\begin{aligned} P_X : \wp(E) &\longrightarrow A \\ A &\longmapsto P(X \in A) \end{aligned}$$
   est une probabilité sur l'espace probabilisable $\big(E, \wp(E)\big)$.

Exercice 6 (tarte à la crème):
L'univers Ω est défini comme $\Omega = [\![1, N]\!]^n$: c'est l'ensemble des n-uplets de $[\![1, N]\!]$. La variable X est la fonction définie comme

$$\begin{aligned} X : \Omega &\longrightarrow [\![1, N]\!] \\ (x_1, \dots, x_n) &\longmapsto \max_{i \in [\![1,n]\!]} x_i. \end{aligned}$$

On cherche la loi de probabilité de X. Attention, on ne cherche pas $P(X = a)$ pour tout a, mais $P(X \leqslant a)$ pour tout a. Cela correspond à l'événement « tous les numérons tirés sont inférieurs à a. » Par équiprobabilité,

$$\begin{aligned} P(X \leqslant a) &= \frac{\mathrm{Card}(X \leqslant a)}{\mathrm{Card}(\Omega)} \\ &= \frac{a^n}{N^n} \end{aligned}$$

Finalement, pour tout $a \in [\![1, N]\!]$, $P(X = a) + P(X \leqslant a - 1) = P(X \leqslant a)$; en effet, $(X \leqslant a - 1) \cup (X = a) = (X \leqslant a)$, et cette union est disjointe. D'où, $\forall a \in [\![1, N]\!]$,

$$\begin{aligned} P(X = a) &= P(X \leqslant a) - P(X \leqslant a - 1) \\ &= \left(\frac{a}{N}\right)^n - \left(\frac{a-1}{N}\right)^n \end{aligned}$$

2 La loi binomiale

Définition 7:
Soient $n \in \mathbb{N}^*$, $p \in]0, 1[$, et $q = 1 - p$. On dit qu'une variable aléatoire discrète X suit une *loi de binomiale* de paramètres (n, p), et on note $X \sim \mathcal{B}(n, p)$ si

$$X(\Omega) = [\![1, n]\!], \quad \text{et} \quad \forall k \in X(\Omega),\ P(X = k) = \binom{n}{k} \cdot p^k \cdot q^{n-k}.$$

Une *épreuve de Bernoulli* de paramètre $p \in]0, 1[$ est une expérience aléatoire qui peut donner deux résultats : un « succès » S avec une probabilité p, ou un « échec » E avec la probabilité $q = 1 - p$.

Proposition 8:
Soient $n \in \mathbb{N}^*$ et $p \in]0, 1[$. Soit X la *vad* égale au nombre de succès parmi n épreuves de Bernoulli de paramètre p. Si ces épreuves sont indépendantes, alors $X \sim \mathcal{B}(n, p)$.

Démonstration:
L'ensemble $(X = k)$ est l'événement « obtenir k succès parmi n essais. » Autrement dit, c'est l'ensemble des n-listes $(x_1, x_2, \dots, x_n)$ dont le nombre de succès vaut k, et chaque $x_i \in \{S, E\}$. Réaliser cet événement, c'est (1) placer k succès parmi les n essais, et il y en a $\binom{n}{k}$ manières ; (2) placer les $n - k$ échecs, il y a 1 manière. Il y a donc $\binom{n}{k}$ listes favorables. La prboabilité de chacune de ces listes est $p^k \times q^{n-k}$ par indépendance.

Exemple 9:
c.f. polycopié

Exercice 10:
On a $X \sim \mathcal{B}(n, p)$. Ainsi, d'après la définition 7, $X(\Omega) = [\![1, n]\!]$, et $\forall k \in X(\Omega)$, $P(X = k) = \binom{n}{k} p^k q^{n-k}$. *Montrons que* $n - X \sim \mathcal{B}(n, q)$ avec $q = 1 - p$. On veut donc montrer que $\forall k \in (n - X)(\Omega)$, $P(n - X = k) = \binom{n}{k} \cdot q^k \cdot p^{n-k}$.

$$\begin{aligned} P(X = n - k) &= \binom{n}{n-k} p^{n-k} q^{n-(n-k)} \\ &= \binom{n}{n-k} q^k p^{n-k} \\ &= \binom{n}{k} q^k p^{n-k} \quad \text{par symétrie des coefficients binomiaux.} \end{aligned}$$

On peut également remarquer que la variable aléatoire $n - X$ est le nombre d'echecs.

3 La loi géométrique

Définition 11:
Soit $p \in]0,1[$. On pose $q = 1-p$. On dit qu'une *vad* T suit une *loi géométrique* de paramètre p, et on note $T \sim \mathcal{G}(p)$ si

$$T(\Omega) = \mathbb{N}^* \quad \text{et} \quad \forall k \in T(\Omega),\ P(T=k) = p \cdot q^{k-1}.$$

On vérifie bien que $\sum_{k\in\mathbb{N}^*} P(X=k) = 1$. En effet, elle vaut

$$\begin{aligned}\sum_{k\in\mathbb{N}^*} pq^{k-1} &= p\sum_{k=1}^{\infty} q^{k-1}\\ &= p\sum_{k=0}^{\infty} q^k\\ &= p \times \frac{1}{1-q}\\ &= \frac{1-q}{1-q} = 1\end{aligned}$$

Proposition 12:
Soit $p \in]0,1[$. Soit T la *vad* égale au temps d'attente du 1$^{\text{er}}$ succès lors d'une suite d'épreuve de Bernoulli de paramètre p. Si ces épreuves sont indépendantes, alors $T \sim \mathcal{G}(p)$.

Démonstration:
L'événement $(X=k)$ vaut $E_1 \cap E_2 \cap \cdots \cap E_{k-1} \cap S_k$, où S_i est l'événement « obtenir un succès au i-ème essai, » et E_i est l'événement « obtenir un échec au i-ème essai.[1] » Par hypothèse d'indépendance, on a

$$P(E_1 \cap E_2 \cap \cdots \cap E_{k-1} \cap S_k) = P(E_1) \times P(E_2) \times \cdots \times P(E_{k-1}) \times P(S_k).$$

Exercice 13 (« Remettre le compteur à zéro »):
Soit X une vad à valeurs dans $\mathbb{N}^*$ telle que

$$\forall k \in \mathbb{N},\ P(X>k) > 0.$$

On dit que X est sans mémoire si

$$\forall k \in \mathbb{N}, \forall n \in \mathbb{N}, \quad P(X > n+k \mid X > k) = P(X > n).$$

On veut montrer que la loi géométrique est une loi sans mémoire, et que c'est la seule.

1. *Montrer que, X est sans mémoire*
   (a) *si, et seulement si $\forall k \in \mathbb{N}, \forall n \in \mathbb{N}$, $P(X > n+k) = P(X>n) \cdot P(X>k)$.*
   (b) *si, et seulement si $\forall k \in \mathbb{N}, \forall n \in \mathbb{N}$, $P(X = n+k \mid X>k) = P(X=n)$.*
2. *On suppose que X suit une loi géométrique de paramètre $p \in]0,1[$.*
   (a) *Calculer la probabilité $P(X>k)$ pour chaque $k \in \mathbb{N}$.*
   (b) *Montrer que X est sans mémoire.*
3. *Réciproquement, montrer que si X est sans mémoire, alors X suit une loi géométrique.*

1. (a)

$$\begin{aligned}X \text{ est sans mémoire} &\overset{\text{def.}}{\iff} \forall k, \forall n,\ P_{(X>k)}(X>n+k) = P(X>n)\\ &\iff P(X>n+k) = P(X>n)\cdot P(X>k)\end{aligned}$$

car

$$P(X>n+k \mid X>k) = \frac{P((X>n+k)\cap(X>k))}{P(X>k)} = \frac{P(X>n+k)}{P(X>k)}$$

comme $((X>n+k)\cap(X>k)) = (X>n+k)$ (par inclusion).

1. Ainsi, on a $S_i = \bar{E}_i$.

(b) On a $(X > n+k-1) = (X = n+k) \cup (X > n+k)$ et cette union est disjointe, d'où

$$P_{(X>k)}(X > n+k-1) = P_{(X>k)}(X = n+k) + P_{(X>k)}(X > n+k).$$

D'où,

$$\begin{aligned} P_{(X>k)}(X = n+k) &= P_{(X>k)}(X > n+k-1) - P(X > n+k) \\ &= P(X > n-1) - P(X > n) \text{ par définition} \\ &= P(X = n) \text{ de même.} \end{aligned}$$

Réciproque à faire.

2. On suppose $X \sim \mathcal{G}(p)$. Ainsi, $X(\Omega) = \mathbb{N}^*$ et $\forall k \in X(\Omega)$, $P(X = k) = p \times q^{k-1}$.

(a) On a $(X > k) = \bigcup_{\ell=k+1}^{\infty} (X = \ell)$, et cette union est disjointe. D'où,

$$\begin{aligned} P(X > k) &= \sum_{\ell=k+1}^{\infty} P(X = \ell) \\ &= \sum_{\ell=k+1}^{\infty} p \times q^{\ell-1} \\ &= p \times q^k \sum_{\ell=0}^{\infty} q^\ell \\ &= p \times q^k \times \frac{1}{1-q} \text{ car } |q| < 1 \\ &= q^k \end{aligned}$$

On en déduit que

$$\forall k \in \mathbb{N},\ P(X > k) = q^k.$$

(b) On utilise le 1.(a) et la question précédente :

$$P(X > n+k) = q^{n+k} = q^n \cdot q^k = P(X > n) \cdot P(X > k).$$

3. On suppose X sans mémoire. On utilise 1.(a) :

$$\forall k,\ \forall n,\ P(X > n+k) = P(X > n) \cdot P(X > k).$$

On cherche une relation de récurrence, on pose donc $k = 1$. Soit $(u_n)_{n\in\mathbb{N}}$ la suite définie par $u_n = P(X > n)$. Ainsi,

$$u_{n+1} = u_n \cdot \underbrace{P(X > 1)}_{\heartsuit}.$$

D'où, par récurrence, $\forall n \in \mathbb{N}$, $u_n = \heartsuit^n \cdot u_0$, et $u_0 = P(X > 0) = 1$ comme $(X > 0) = \Omega$. Ainsi,

$$\forall n \in \mathbb{N},\ u_n = \heartsuit^n = P(X > n).$$

On a, pour $n \in \mathbb{N}^*$, $(X > n-1) = (X = 1) \cup (X > n) = (X = n) \cup P(X > n)$, et cette union est disjointe. D'où, $\heartsuit^{n-1} = P(X = n)/\heartsuit^{n-1} - \heartsuit^n$.

4 La loi de Poisson

Définition 14:
Soit un réel $\lambda > 0$. On dit qu'une variable aléaroire discrète X suit *une loi de Poisson* de paramètre λ, et on note $X \sim \mathcal{P}(\lambda)$ si

$$X(\Omega) = \mathbb{N} \quad \text{et} \quad \forall k \in X(\Omega),\ P(X = k) = \mathrm{e}^{-\lambda} \cdot \frac{\lambda^k}{k!}.$$

On vérifie que $\sum_{k\in X(\Omega)} P(X=k)=1$. En effet,

$$\begin{aligned}\sum_{k\in X(\Omega)} P(X=k) &= \sum_{k=0}^{\infty} \mathrm{e}^{-\lambda}\frac{\lambda^k}{k!}\\ &= \mathrm{e}^{-\lambda}\sum_{k=0}^{\infty}\frac{\lambda^k}{k!}\\ &= \mathrm{e}^{-\lambda}\cdot \mathrm{e}^{\lambda} = 1\end{aligned}$$

Proposition 15:
Soit un réel $\lambda > 0$, et pour $n \in \mathbb{N}^*$, soit X_n une variable aléatoire suivant la loi binomiale $\mathcal{B}(n, p_n)$ avec $p_n \in \left]0,1\right[$. Si $n\cdot p_n \xrightarrow[n\to\infty]{} \lambda$, alors

$$\forall k\in\mathbb{N},\ P(X_n=k)\xrightarrow[n\to\infty]{} \mathrm{e}^{-\lambda}\cdot\frac{\lambda^k}{k!}.$$

Démonstration:
On pose, $q_n = 1-p_n$. On a

$$\binom{n}{k} p_n^k\,(1-p_n)^{n-k} = \frac{n!}{k!\cdot(n-k)!}\cdot p_n^k\cdot(1-p_n)^{n-k}..$$

Or,

$$\begin{aligned}p_n^k\cdot\frac{n!}{(n-k)!} &= p_n^k\cdot n\times(n-1)\times\cdots\times(n-k+1)\\ &= p_n^k\cdot n\times n\left(1-\frac{1}{n}\right)\cdots n\left(1-\frac{k-1}{n}\right)\\ &= (p_n n)^k\cdot\underbrace{\left(1-\frac{1}{k}\right)\times\cdots\times\left(1-\frac{k-1}{n}\right)}_{\xrightarrow[n\to\infty]{}1}\\ &\to \lambda^k.\end{aligned}$$

De plus, $(1-p_n)^{n-k} = \mathrm{e}^{(n-k)\ln(1-p_n)}$, et

$$\begin{aligned}(n-k)\ln(1-p_n) &= (n-k)\big(-p_n+o(p_n)\big)\\ &= -np_n + o(np_n)\\ &\to -\lambda.\end{aligned}$$

Par continuité de l'exponentielle, on a $(1-p_n)^{n-k}\xrightarrow[n\to\infty]{}\mathrm{e}^{-\lambda}$. Ainsi,

$$\forall k\in\mathbb{N},\ P(X_n=k)\xrightarrow[n\to\infty]{} \mathrm{e}^{-\lambda}\cdot\frac{\lambda^k}{k!}.$$

Remarque 16:
c.f. polycopié

5 Espérance

Définition 17:
Soient $(\Omega,\mathcal{A},P)$ un espace probabilisé, et X une variable aléatoire réelle discrète (*vard*) telle que $X(\Omega)=\{a_i \mid i\in I\}\subset\mathbb{R}$. Si la série $\sum a_i\,P(X=a_i)$ converge absolument[2], alors

1. on dit que X est *d'espérance finie* ou que X *possède une espérance*, ou que $X\in L^1$;
2. cette *espérance*, notée $\mathbb{E}(X)$ est le nombre réel

$$\mathbb{E}(X)=\sum_{i\in I} a_i\,P(X=a_i).$$

2. Ainsi, la série devient une famille sommable, et la somme devient commutative, même si elle est infinie.

Remarque 18: 1. La valeur de l'espérance ne dépend pas de l'ordre des valeurs a_i, c'est pour cela que l'on demande la convergence absolue.

2. Une *vard* peut ne pas avoir une espérance finie, par exemple $p_n = \frac{6}{\pi^2 n^2}$. On a bien $\sum_{n=1}^{\infty} p_n = 1$ mais la série $\sum n\, p_n$ diverge.
3. Si $X(\Omega)$ est fini, alors X est nécessairement d'espérance finie, car $\mathbb{E}(X)$ est une somme finie.
4. Si X est bornée [3](vue comme une fonction), alors X est d'espérance finie.

   Démonstration:
   En effet, $\forall a \in X(\Omega)$, $0 \leqslant |a\, P(X = a)| \leqslant M\, P(X = a)$ d'où la série $\sum |a\, P(X = a)|$ converge, car $\sum M\, P(X = a) = M \sum P(X = a)$ et $\sum P(X = a)$ converge.
5. L'espérance est linéaire : si X est un *vard* d'espérance finie, alors $\alpha X + \beta$ aussi, et

$$\mathbb{E}(\alpha X + \beta) = \alpha\, \mathbb{E}(X) + \beta.$$

   De plus, l'application

$$\begin{aligned} \mathbb{E} : \text{ensemble des } \textit{vard} &\longrightarrow \mathbb{R} \\ X &\longmapsto \mathbb{E}(X) \end{aligned}$$

   est une forme linéaire.
6. Si X est d'espérance finie, alors $|X|$ aussi, et $\left|\mathbb{E}(X)\right| \leqslant \mathbb{E}\left(|X|\right)$ (par inégalité triangulaire). En particulier, si X est positive, alors son espérance est positive.

Exercice 19:
Montrer que

1. *si $X \sim \mathcal{B}(n, p)$, alors X est d'espérance finie, et $\mathbb{E}(X) = n \cdot p$. Indication, utiliser la formule*

$$\forall k \in \mathbb{N}^*,\ \forall n \geqslant k,\ k\binom{n}{k} = n\binom{n-1}{k-1}.$$

2. *si $T \sim \mathcal{G}(p)$, alors T est d'espérance finie, et $\mathbb{E}(T) = \frac{1}{p}$.*
3. *si $X \sim \mathcal{P}(\lambda)$, alors X est d'espérance finie, et $\mathbb{E}(X) = \lambda$.*

1. La variable aléatoire X représente le nombre de succès, de n épreuves de Bernoulli indépendantes, et la probabilité d'un succès est p. De plus, on a $X(\Omega) = [\![0, n]\!]$, qui est un ensemble fini, il possède une espérance. On calcule

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

2. La variable T correspond au temps d'attente du 1er succès, sachant que la probabilité d'un succès est p. Ainsi, $T(\Omega) = \mathbb{N}^*$, et $\forall x \in T(\Omega)$, $P(T = k) = p \cdot q^{k-1}$. La variable aléatoire est d'espérance finie car la série $\sum k\, P(T = k)$ converge absolument :

$$\left|k\, P(T = k)\right| = k\, P(T = k) = k\, p\, q^{k-1}.$$

3. *i.e.* $\exists M \in \mathbb{R}^+$, $\forall \omega \in \Omega$ $|X(\omega)| \leqslant M$.

D'où $\sum |k\,P(T=k)| = \sum k\,p\,q^{k-1} = p\sum k\,q^{k-1}$. Or, la série entière $\sum x^k$ a pour rayon de converge 1. Et, on peut dériver terme à terme une série entière sans changer son rayon de convergence. D'où, le rayon de convergence de $\sum k\,x^{k-1}$ est aussi égal à 1. Or, $q \in]-1,1[$, d'où la série $\sum k\,q^{k-1}$ converge.
De plus,

$$\forall x \in]-1,1[, \quad \sum_{k=0}^{\infty} x^k = \frac{1}{1-x}.$$

Et, on peut dériver terme à terme une série entière sans changer son rayon de convergence. D'où,

$$\forall x \in]-1,1[, \quad \sum_{k=1}^{\infty} kx^{k-1} = \frac{\mathrm{d}}{\mathrm{d}x}\frac{1}{1-x} = \frac{1}{(1-x)^2}.$$

D'où,

$$\mathbb{E}(T) = \sum_{k=1}^{\infty} k\,P(T=k) = p \times \frac{1}{(1-q)^2} = \frac{1}{1-q} = \frac{1}{p}.$$

3. On a $X(\Omega) = \mathbb{N}$, et $\forall k \in X(\Omega)$, $P(X=k) = \mathrm{e}^{-k}\,\frac{\lambda^k}{k!}$. La *vard* X possède une espérance car la série $\sum k\,P(X=k)$ converge absolument. En effet, $\sum \left|k\,\mathrm{e}^{-\lambda}\,\frac{\lambda^k}{k!}\right| = \lambda \mathrm{e}^{-\lambda}\sum k\,\frac{\lambda^{k-1}}{k!}$. Or, la série entière $\sum x\,\frac{k}{k!}$ a pour rayon de converge $+\infty$. Et, on peut dériver terme à terme une série entière sans changer son rayon de convergence. D'où, le rayon de convergence de $\sum k\,\frac{x^{k-1}}{k!}$ est aussi égal à $+\infty$. De plus,

$$\begin{aligned}
\mathbb{E}(X) &= \sum_{k\in X(\Omega)} k\,P(X=k) \\
&= \sum_{k=0}^{\infty} k\,\mathrm{e}^{-\lambda}\,\frac{\lambda^k}{k!} \\
&= \sum_{k=1}^{\infty} k\,\mathrm{e}^{-\lambda}\,\frac{\lambda^k}{k!} \\
&= \mathrm{e}^{-\lambda}\,\lambda \sum_{k=1}^{\infty} k\,\frac{\lambda^{k-1}}{k!} \\
&= \lambda\,\mathrm{e}^{-\lambda}\sum_{k=0}^{\infty}\frac{\lambda^k}{k!} \\
&= \lambda\,\mathrm{e}^{-\lambda}\,\mathrm{e}^{\lambda} = \lambda
\end{aligned}$$

PROPOSITION 20 (version non rigoureuse):

$$\begin{aligned}
\mathbb{E}(X) &= \sum_{k=0}^{\infty} k\,P(X=k) \\
&= 0\times P(X=0) + 1\times P(X=1) + 2\times P(X=2) + 3\times P(X=3) + \cdots \\
&= \quad P(X=1) \\
&\quad + P(X=2) + P(X=2) \\
&\quad + P(X=3) + P(X=3) + P(X=3) \\
&\quad + \quad \vdots \quad + \quad \vdots \quad + \quad \vdots \quad + \quad \ddots \\
&= \quad P(X \geqslant 1) + P(X \geqslant 2) + P(X \geqslant 3) + \quad \cdots
\end{aligned}$$

Les hypothèse de ce théorème sont

- la variable X est d'espérance finie.
- la série $\sum P(X \geqslant n)$ converge (c'est donc une famille sommable, car $P(X \geqslant n) \geqslant 0$).

Ainsi, on peut sommer par paquets.

REMARQUE 21:
Si X est une *vard* est $\varphi : X(\Omega) \to \mathbb{R}$, alors

1. $\varphi \circ X$ est aussi une *vard* notée $\varphi(X)$;
2. on pose $(\varphi \circ X)(\Omega) = \{b_j \mid j \in J\}$. Si $\varphi(X)$ possède une espérance $\mathbb{E}\big(\varphi(X)\big)$, alors cette espérance est égale à
$$\sum_{j \in J} b_j \; P\big(\varphi(X) = b_j\big)$$
par définition de l'espérance, mais aussi à
$$\sum_{i \in I} \varphi(a_i) \; P(X = a_i),$$
d'après le théorème suivant, que nous admettrons.

Théorème 22:
Soient $\Omega, \mathcal{A}, P$ un espace probabilisé, X une *vard*, et $\varphi : X(\Omega) \to \mathbb{R}$. On pose $X(\Omega) = \{a_i \mid i \in I\}$. Alors, $\varphi(X)$ est d'espérance finie si, et seulement si la série $\sum \varphi(a_i)\, P(X = a_i)$ converge absolument. Et, alors, $\mathbb{E}(\varphi(X))$ vaut la somme $\sum_{i \in I} \varphi(a_i)\, P(X = a_i)$ de cette série.

6 Variance et écart-type

Plus tard (définition 25), on définira respectivement la *variance* et l'*écart-type* comme
$$\mathrm{V}(X) = \mathbb{E}\Big(\big[X - E(X)\big]^2\Big) \qquad \sigma(X) = \sqrt{\mathrm{V}(X)}.$$

Définition 23:
On définit le *moment d'ordre* $k \in \mathbb{N}$ comme
$$\sum_{a \in X(\Omega)} a^k \; P(X = a).$$
On dit qu'une variable aléatoire *possède un moment d'ordre* k si la série $\sum {a_n}^k P(X = a_n)$ converge absolument.

Ainsi, le moment d'ordre 1 est $\mathbb{E}(X)$, le moment d'ordre 2 est $\mathbb{E}(X^2)$, le moment d'ordre 3 est $\mathbb{E}(X^3)$, etc.

Lemme 24:
Si une *vard* possède un moment d'ordre $k+1$, alors elle possède aussi un moment d'ordre k.

Démonstration:
Pour tout réel, $x \geqslant 0$, alors $0 \leqslant x^k \leqslant x^{k+1} + 1$ (distinguer le cas $x \geqslant 1$ et $x < 1$). Ainsi, en multipliant par $P(X = a)$, on a
$$x^k \; P(X = a) \leqslant x^{k+1} \; P(X = a) + P(X = a).$$
La série $\sum P(X = a)$ converge absolument, et la série $\sum a^{k+1} \; P(X = a)$ converge absolument aussi. Alors, la série $\sum a^k \; P(X = a)$ converge absolument.

Proposition – Définition 25:
Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé, et soit X une *vard*. Si X^2 est d'espérance finie, alors X aussi, et on appelle *variance* le réel positif
$$\mathrm{V}(X) = \mathbb{E}\Big(\big[X - \mathbb{E}(X)\big]^2\Big) = \underbrace{\mathbb{E}(X^2) - \big(\mathbb{E}(X)\big)^2}_{\text{Relation de König \& Huygens}} \geqslant 0.$$
L'*écart-type* $\sigma(X)$ est la racine carrée de la variance :
$$\sigma(X) = \sqrt{\mathrm{V}(X)}.$$

Démonstration:
On pose $\mu = \mathbb{E}(X)$, et on a $[X - \mu]^2 = X^2 - 2\mu X + \mu^2$. D'où, par linéarité de l'espérance,
$$\begin{aligned} \mathbb{E}\big((X-\mu)^2\big) &= \mathbb{E}(X^2 - 2\mu X + \mu^2) \\ &= \mathbb{E}(X^2) - 2\mu\mathbb{E}(X) + \mu^2 \\ &= \mathbb{E}(X^2) - 2\mu^2 + \mu^2 \\ &= \mathbb{E}(X^2) - \big(\mathbb{E}(X)\big)^2. \end{aligned}$$

De plus, d'après le lemme précédent, si X^2 est d'espérance finie, alors X est d'espérance finie.

Remarque 26: 1. La variance mesure la *dispersion*, ou l'*étalement* des valeurs a_i autour de l'espérance $\mathbb{E}(X)$. En particulier, s'il existe $a \in \mathbb{R}$ tel que $P(X = a) = 1$, alors $\mathbb{E}(X) = a$ et $\mathbb{V}(X) = 0$. (C'est même une équivalence.)

2. Si la variable X a une unité (km/s, V/m, etc.), alors l'écart type a la même unité (d'où l'intérêt de calculer la racine carrée de la variance).
3. Soient α et β deux réels. Si X^2 est d'espérance finie, alors

$$\mathbb{V}(\alpha X + \beta) = \alpha^2 \cdot \mathbb{V}(X).$$

(Une translation ne change pas la dispersion des valeurs, et multiplier par un réel multiplie l'espérance, mais aussi la dispersion, d'où le carré.)

Exercice 27:
Montrer que

1. *si $X \sim \mathcal{B}(n, p)$, alors X^2 est d'espérance finie et $\mathbb{V}(X) = n \cdot p \cdot q$.*
2. *si $T \sim \mathcal{G}(p)$, alors T^2 est d'espérance finie et $\mathbb{V}(T) = \frac{q}{p^2}$.*
3. *si $X \sim \mathcal{P}(\lambda)$, alors X^2 est d'espérance finie et $\mathbb{V}(X) = \lambda$.*

1. Si $X \sim \mathcal{B}(n, p)$, alors $X(\Omega) = [\![0, n]\!]$ et, pour $k \in X(\Omega)$, $P(X = k) = \binom{n}{k} p^k q^{n-k}$. On a déjà montré que $\mathbb{E}(X) = n \cdot p$. On va montrer que $\mathbb{V}(X) = n\,p\,q$. La variable aléatoire X^2 est d'espérance finie car $X(\Omega)$ est fini. Et,

$$\begin{aligned}\mathbb{E}(X^2) &= \sum_{k=0}^{n} k^2\, P(X = k)\\ &= \sum_{k=0}^{n} k^2 \binom{n}{p} p^k q^{n-k}\\ &= \dots\end{aligned}$$

En effet, d'après la "petite formule," on a

$$\forall k \geqslant 1, \quad k\binom{n}{k} = n\binom{n-1}{k-1}$$

d'où, $\binom{(k-1)\binom{n-1}{k-1} = (n-1)}{\binom{n-2}{k-2}}$. Ainsi,

$$\forall k \geqslant 2, \quad k(k-1)\binom{n}{k} = n(n+1)\binom{n-2}{k-2}.$$

2. Si $T \sim \mathcal{G}(p)$, alors $T(\Omega) = \mathbb{N}^*$ et $\forall k \in T(\Omega)$, $P(T = k) = p \times q^{k-1}$. On a déjà prouvé que $\mathbb{E}(T) = \frac{1}{p}$. On veut montrer que $\mathbb{V}(T) = \frac{q}{p^2}$. Montrons que la variable T^2 possède une espérance : la série $\sum k^2\, P(T = k)$ converge absolument car $k^2\, P(T = k) = k^2 \cdot p \cdot q^{k-1}$. Or, pour $k \geqslant 2$, $\frac{\mathrm{d}^2}{\mathrm{d}x^2} x^k = k(k-1)\, x^{k-2}$. Et, on peut dériver terme à terme une série entière sans changer son rayon de convergence, et la série $\sum x^k$ a pour rayon de convergence 1. D'où, $\sum k(k-1)\, x^{k-2}$ a pour rayon de convergence 1. Or, $q \in \,]0, 1[\, \subset \,]-1, 1[$ donc la série $\sum k(k-1) q^{k-2}$ converge. De plus, $\sum k(k-1)\, q^{k-2} = \sum k^2\, q^{k-2} - \sum k\, q^{k-2}$. D'où, $\sum k^2 q^{k-2} = \sum k(k-1)\, q^{k-2} + \sum k\, q^{k-2}$, qui converge.

Par suite,

$$\begin{aligned}
\sum_{k=1}^{\infty} k^2\, P(T=k) &= \sum_{k=1}^{\infty} k^2\, p\, q^{k-1} \\
&= p + pq \sum_{k=2}^{\infty} k^2 q^{k-2} \\
&= p + pq \sum_{k=2}^{\infty} k(k-1)\, q^{k-2} + p \sum_{k=2}^{\infty} k\, q^{k-1} \\
&= p + pq\, \frac{2}{(1-q)^3} + p\left(\frac{1}{(1-q)^2} - 1\right) \qquad \textit{c.f.} \text{ en effet après} \\
&= p + pq\, \frac{2}{p^3} + p\left(\frac{1}{p^2} - 1\right) \\
&= \frac{2q}{p^2} + \frac{1}{p} \\
&= \frac{2q+p}{p^2} \\
&= \frac{2q+(1-q)}{p^2} \\
&= \frac{q+1}{p^2}.
\end{aligned}$$

En effet, $\forall x \in\ ]-1,1[$, $\sum_{k=0}^{\infty} x^k = \frac{1}{1-x}$. D'où, pour $x \in\ ]-1,1[$,

$$\sum_{k=1}^{\infty} k\, x^{k-1} = \frac{1}{(1-x)^2} \quad \text{et} \quad \sum_{k=2}^{\infty} k(k-1)\, x^{k-2} = \frac{2}{(1-x)^3}.$$

Ainsi, $\mathbb{E}(T^2) = \frac{q+1}{p^2}$. D'où

$$\begin{aligned}
\mathbb{V}(T) &= \mathbb{E}(T^2) - \big(\mathbb{E}(T)\big)^2 \\
&= \frac{q+1}{p^2} - \left(\frac{1}{p}\right)^2 \\
&= \frac{q}{p^2}
\end{aligned}$$

3. À tenter

7 Les inégalités de Markov et de Bienaymé–Tchebychev, inégalités de concentration

Lemme 28 (Markov):
Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé, et soit X une variable aléatoire positive. Si X est d'espérance finie, alors

$$\forall a > 0, \quad P(X \geqslant a) \leqslant \frac{\mathbb{E}(X)}{a}.$$

Démonstration:
On suppose X d'espérance finie. Ainsi, on a

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

Soit I l'ensemble $I = \{x \in X(\Omega) \mid x \geqslant a\}$. Alors,

$$\mathbb{E}(X) = \underbrace{\sum_{x \in I} x\, P(X = x)}_{\text{ici } x \geqslant a} + \underbrace{\sum_{x \in X(\Omega) \setminus I} x\, P(X = x)}_{\geqslant 0 \text{ par hypothèse}}.$$

D'où,

$$\mathbb{E}(X) \geqslant \sum_{x \in I} x\, P(X = x) \geqslant \sum_{x \in I} a\, P(X = x) = a \sum_{x \in I} P(X = x) \geqslant a\, P(x \geqslant a).$$

Proposition 29 (Bienaymé–Tchebychev):
Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé, et soit X une *vard*. Si X^2 est d'espérance finie, alors

$$\forall a > 0, \qquad P\Big(\big|X - \mathbb{E}(X)\big| \geqslant a\Big) \leqslant \frac{\mathrm{V}(X)}{a^2}.$$

Démonstration:
On pose $\mu = \mathbb{E}(X)$. L'événement $\big(|X - \mu| \geqslant a\big) = \big((X - \mu)^2 \geqslant a^2\big)$, d'où, les probabilités

$$P\big(|X - \mu| \geqslant a\big) = P\big(\underbrace{(X - \mu)^2}_{\geqslant 0} \geqslant \underbrace{a^2}_{} \geqslant 0\big).$$

On valide donc *une* des hypothèses de l'inégalité de Markov. De plus, l'autre hypothèse est vérifiée : X^2 est d'espérance finie, donc $(X - \mu)^2$ aussi. On en déduit, d'après le lemme de Markov, que

$$P\big((X - \mu)^2 \geqslant a^2\big) \leqslant \frac{\mathbb{E}\big((X - \mu)^2\big)}{a^2} = \frac{\mathrm{V}(X)}{a^2}.$$

8 Série génératrice

Définition 30:
Soit X une *vad* telle que $X(\Omega) \subset \mathbb{N}$. La *série génératrice* de X est la série entière $\sum a_n x^n$ de coefficients $a_n = P(X = n)$.

La série $\sum a_n$ converge car sa somme vaut $\sum_{n=0}^{\infty} a_n = 1$. D'où,

- le rayon de convergence R de la série est supérieur ou égal à 1.
- la série génératrice converge normalement sur $[-1, 1]$, car la série $\sum |a_n|$ converge, or, $\forall x \in [-1, 1]$, $|p_n t^n| \leqslant |p_n|$, d'où la convergence normale. D'où la *fonction génératrice*

  $$\mathrm{G}_X : t \longmapsto \sum_{n=0}^{\infty} p_n t^n$$

  est définie et même continue sur $[-1, 1]$, car la convergence est uniforme.
- la fonction génératrice G_X est de classe $\mathcal{C}^\infty$ sur $]-1, 1[$ et

  $$\forall k \in \mathbb{N}, \quad P(X = k) = a_n \frac{{\mathrm{G}_X}^{(k)}(0)}{k!}.$$

  La fonction génératrice de X permet donc de retrouver la loi de probabilité de X.

Cette série génératrice permet de calculer l'espérance et la variance de X. Si $R > 1$, alors $1 \in\,]-R, R[$, d'où

$$\mathrm{G}'_X(1) = \sum_{n=1}^{\infty} n a_n \qquad \text{et} \qquad \mathrm{G}''_X(1) = \sum_{n=2}^{\infty} n(n-1)\, a_n,$$

car on peut dériver terme à terme sans changer de convergence. On en déduit que X^2 est d'espérance finie et que

$$\mathbb{E}(X) = \mathrm{G}'_X(1) \qquad \text{et} \qquad \mathrm{V}(X) = G''_X(1) + G'_X(1) - \big[G'_X(1)\big]^2.$$

Également, même si inintéressant du point de vue théorique, ceci peut être utile pour vérifier les résultats en exercice

$$\mathrm{G}_X(1) = \sum_{n=0}^{\infty} a_n = \sum_{n=0}^{\infty} P(X = n) = 1.$$

PROPOSITION 31:
Soit X une *vard* telle que $X(\Omega) \subset \mathbb{N}$, et soit G_X sa fonction génératrice.

1. X est d'espérance finie si, et seulement si la fonction G_X est dérivable en 1. Dans ce cas,
$$\mathbb{E}(X) = \mathrm{G}'_X(1).$$
2. X^2 est d'espérance finie si, et seulement si la fonction G_X est deux fois dérivable en 1. Dans ce cas,
$$\mathrm{V}(X) = \mathrm{G}''_X(1) + \mathrm{G}'_X(1) - \left[\mathrm{G}'_X(1)\right]^2.$$

DÉMONSTRATION:
On remarque que $\mathrm{G}''_X(1) = \sum_{n=2}^{\infty} n(n-1)\, a_n = \mathbb{E}\big(X(X-1)\big) = \mathbb{E}(X^2) - \mathbb{E}(X)$. Ainsi,

$$\begin{aligned}
\mathrm{V}(X) &= \mathbb{E}\left[\big(X - \mathbb{E}(X)\big)^2\right] \\
&= \mathbb{E}(X^2) - \big(\mathbb{E}(X)\big)^2 \\
&= \mathbb{E}(X^2) - \mathbb{E}(X) + \mathbb{E}(X) - \big(\mathbb{E}(X)\big)^2 \\
&= \mathrm{G}''_X(1) + \mathrm{G}'_X(1) - \big(\mathrm{G}'_X(1)\big)^2
\end{aligned}$$

Mais, la fonction G_X est elle,

— dérivable en 1 ? Oui, si la variable X est d'espérance finie.
— dérivable deux fois en 1 ? Oui, si la variable X^2 est d'espérance finie.

Le programme dit que l'on doit être capable de le retrouver rapidement.

EXERCICE 32:
Soient $p \in \,]0, 1[$, et $\lambda > 0$. On pose $q = 1 - p$. Soit X une variable aléatoire. Montrer que

1. *si $X \sim \mathcal{B}(n, p)$, alors* $\quad \forall t \in \mathbb{R}, \quad \mathrm{G}_X(t) = (pt + q)^n$;
2. *si $T \sim \mathcal{G}(p)$, alors* $\quad \forall t \in \left]-\frac{1}{q}, \frac{1}{q}\right[\quad \mathrm{G}_T(t) = \frac{pt}{1-qt}$;
3. *si $X \sim \mathcal{P}(\lambda)$, alors* $\quad \forall t \in \mathbb{R}, \quad \mathrm{G}_X(t) = \mathrm{e}^{-\lambda} \cdot \mathrm{e}^{\lambda t}$.

En déduire l'espérance et la variance de chacune de ces vard.

1. La série génératrice est $\sum P(X = k)\, t^k$, et la fonction génératrice est $\sum_{k \in X(\Omega)} P(X = k)\, t^k$. Comme $X \sim \mathcal{B}(n, p)$, on a $X(\Omega) = [\![0, n]\!]$, et $\forall k \in X(\Omega)$, $P(X = k) = \binom{n}{k} p^k q^{n-k}$. La série ne peut pas diverger, car il y a un nombre fini de termes. D'où,
$$\forall t \in \,]-\infty, +\infty[, \quad \begin{aligned}[t] \mathrm{G}_X(t) &= \sum_{k=0}^{n} \binom{n}{k} q^{n-k}\, t^k \\ &= (pt + q)^n \end{aligned}$$

   La fonction G_X est dérivable en 1, donc la variable aléatoire X est d'espérance finie, et $\mathbb{E}(X) = \mathrm{G}'_X(1)$. Or, $\forall t \in \mathbb{R}$, $\mathrm{G}'_X(t) = n\, p\, (pt + q)^{n-1}$. D'où, $\mathbb{E}(X) = \mathrm{G}'_X(1) = n\, p$. Mieux : G_X est deux fois dérivable en 1. Ainsi, X^2 est d'espérance finie, et
$$\begin{aligned}
\mathrm{V}(X) &= \mathrm{G}''_X(1) + \mathrm{G}'_X(1) - \left[\mathrm{G}'_X(1)\right]^2 \\
&= n \cdot (n-1) \cdot p^2 + n \cdot p - (n \cdot p)^2 \\
&= n \cdot p - n \cdot p^2 = n \cdot p \cdot (1 - p) \\
&= n \cdot p \cdot q
\end{aligned}$$

2. Si $T \sim \mathcal{G}(p)$, alors $T(\Omega) = \mathbb{N}^*$, et $\forall k \in T(\Omega)$, $P(T = k) = p \times q^{k-1}$. La série génératrice de la variable T est

$$\sum P(T = k)t^k = \sum p\, q^{k-1}\, t^k = p\, t \sum (qt)^{k-1},$$

c'est une série géométrique de raison qt. Elle converge si, et seulement si $|qt| < 1$. D'où, le rayon de convergence de la série génératrice vaut $R = \frac{1}{q}$. Et,

$$\begin{aligned} \forall t \in \left]-\frac{1}{q}, \frac{1}{q}\right[, \quad \mathrm{G}_T(t) &= p\, t \sum_{k=1}^{\infty} (qt)^{k-1} \\ &= pt \sum_{k=0}^{\infty} (qt)^k \\ &= pt \cdot \frac{1}{1 - qt} \text{ car } |qt| < 1 \end{aligned}$$

G_X est dérivable en 1, d'où, la variable aléatoire T est d'espérance finie, et $\mathrm{E}(T) = \mathrm{G}'_T(1)$.

$$\forall t \in \left]-\frac{1}{q}, \frac{1}{q}\right[, \quad \mathrm{G}'_T(t) = \frac{p(1 - qt) - pt(-q)}{(1 - qt)^2} = \frac{p}{(1 - qt)^2}.$$

D'où, $\mathrm{E}(t) = \frac{p}{(1-q)^2} = \frac{1}{p}$.