Variables aléatoires discrètes

I Généralités – Définitions et propriétés

I.1 Définition générale d'une variable aléatoire

Définition

Soient $(\Omega, \mathcal{A})$ et $(\Omega', \mathcal{A}')$ deux espaces probabilisables et $X : \Omega \longrightarrow \Omega'$. On dit que X est une variable aléatoire si l'image réciproque de tout événement par X est un événement. Autrement dit :

$$X \text{ est une variable aléatoire } \iff \forall B \in \mathcal{A}',\ X^{-1}(B) \in \mathcal{A}$$

Notation : Dans le cadre des variables aléatoires, l'événement $X^{-1}(B)$ est toujours noté $[X \in B]$.

Nous traduisons cette définition dans le cas particulier du programme :

I.2 Variables aléatoires discrètes

Définition

Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé, soit E un ensemble, soit $X : \Omega \longrightarrow E$ une application. On dit que X est une variable aléatoire discrète lorsque :

i. $X(\Omega)$ est fini ou dénombrable

ii. $\forall y \in X(\Omega)$, $X^{-1}(\{y\})$ est un événement

Notation : l'événement $X^{-1}(\{y\})$ est toujours noté $[X = y]$. Cette définition coincide avec la précédente en effet on a :

Proposition

Soit $X : \Omega \longrightarrow E$ une application telle que $X(\Omega)$ est fini ou dénombrable. Alors les deux propositions suivantes sont équivalentes :

i. X est une variable aléatoire discrète

ii. $\forall B \subset E$, l'ensemble $[X \in B]$ est un événement

Démonstration.

$\boxed{\Leftarrow}$ Il suffit de prendre $B = \{y\}$ où $y \in X(\Omega)$.

$\boxed{\Rightarrow}$ $[X \in B] = [X \in B \cap X(\Omega)] = \cup_{b \in B \cap X(\Omega)} [X = b]$. Or $[B \cap X(\Omega)]$ est dénombrable. $[X \in B]$ est une union dénombrable d'événements, c'est donc un événement.

Remarque. Si Ω est fini, toutes les applications définies sur Ω sont des variables aléatoires. Ce résultat est toujours valable si Ω est dénombrable et muni de la tribu $\mathcal{A} = \mathcal{P}(\Omega)$.

I.3 Premiers exemples

Indicatrices :

Soit $A \subset \Omega$ un événement. $X = \mathbb{1}_A$ est une variable aléatoire discrète.

Technique de démonstration sur un exemple

On joue à pile ou face indéfiniment. On note X l'indice du premier pile, avec éventuellement $X = +\infty$. Montrer que X est une variable aléatoire réelle discrète.

I.4 Propriétés

Proposition

Soit $X : (\Omega, \mathcal{A}, \mathbb{P}) \longrightarrow E$ une variable aléatoire discrète. Soit $f : E \longrightarrow F$ une application quelconque. Alors $f \circ X$ est une variable aléatoire discrète.

Proposition (vecteur aléatoire)

On se donne, pour $i \in [\![1, n]\!]$, $X_i : (\Omega, \mathcal{A}, \mathbb{P}) \longrightarrow E_i$ une variable aléatoire discrète. Posons :

$$X : \left| \begin{array}{lcl} (\Omega, \mathcal{A}, \mathbb{P}) & \longrightarrow & E_1 \times \cdots \times E_n \\ \omega & \longmapsto & (X_1(\omega), ..., X_n(\omega)) \end{array} \right.$$

Alors X est une variable aléatoire discrète.

Démonstration. $X(\Omega)$ est dénombrable, et comme chaque X_i est un variable aléatoire, $[X = (x_1, ..., x_n)]$ est l'intersection des $[X_i = x_i]$, c'est donc un événement.

Proposition

Si X et Y sont deux variables aléatoires discrètes sur $(\Omega, \mathcal{A}, \mathbb{P})$, si f est une fonction quelconque, alors $Z = f(X, Y)$ est une variable aléatoire discrète.

En particulier, $X + Y$, $X \times Y$, λX....... sont des variables aléatoires discrètes.

Ces propositions prouvent que la notion de variable aléatoire est très souple. Voici un exemple un peu plus sophistiqué pour s'en rendre compte.

Si X et Y sont deux variables aléatoires discrètes à valeurs réelles, posons :

$$Z = \begin{cases} XY \text{ si } e^X \geqslant Y \\ X - 1 \text{ sinon} \end{cases}$$

Alors Z est une variable aléatoire (en effet, on peut par exemple écrire $Z = XY \mathbb{1}_{e^X \geqslant Y} + (X-1)\mathbb{1}_{e^X < Y}$).

Dans la pratique, toutes les applications que nous rencontrerons seront des variables aléatoires.

I.5 Loi d'une variable aléatoire

Définition

Soit $X = (\Omega, \mathcal{A}, \mathbb{P}) \longrightarrow E$ une variable aléatoire discrète. Alors en posant, pour $B \subset X(\Omega)$, $\mathbb{P}_X(B) = \mathbb{P}(\{X \in B\})$, on définit une loi de probabilité sur $X(\Omega)$ muni de la tribu $\mathcal{P}(X(\Omega))$. $\mathbb{P}_X$ s'appelle la loi de X.

Démonstration.

Proposition

La loi d'une variable aléatoire discrète est entièrement déterminée par la famille des réels $p_x = \mathbb{P}(X = x)$ pour x décrivant $X(\Omega)$ (les p_x constituent la distribution de probabilités)

Dans la pratique, pour déterminer la loi d'une variable aléatoire :

-On commence par identifier l'ensemble $X(\omega)$
-On détermine pour tout $x \in X(\omega)$ la probabilité $\mathbb{P}(X = x)$.

Exemple. On joue à pile ou face. On note p la probabilité que la pièce tombe sur pile, et $q = 1 - p$. On note X la variable aléatoire égale au rang du premier pile. Alors :

$X(\Omega) = \mathbb{N}^* \cup \{+\infty\}$,
pour tout $k \in \mathbb{N}^*$, $\mathbb{P}(X = k) = pq^{n-1}$
$\mathbb{P}(X = +\infty) = 0$.

La loi de X est enitèrement déterminée.

Variables identiquement distribuées
Deux variables aléatoires différentes peuvent avoir la même loi. Par exemple, pour $1 \leq i \leq 6$ notons X_i la variable aléatoire indicatrice du résultat d'un lancer de dé. La loi de X_i est donnée par :

Ainsi toutes les variables X_i ont la même loi, mais ces variables sont tout à fait différentes.

Notation : Lorsque deux variables X et Y ont la même loi, on notera $X \sim Y$.

Des variables de même loi sont dites identiquement distribuées.

II Espérance d'une variable aléatoire réelle

II.1 Définition

Définition (cas des variables aléatoires réelles discrètes positives)

Soit $X : (\Omega, \mathcal{A}, \mathbb{P}) \longrightarrow \mathbb{R}_+$ une variable aléatoire réelle discrète **positive**. On pose dans $[0, = \infty]$:

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

$E(X)$ s'appelle l'espérance de X

Remarque.
$X(\Omega)$ est dénombrable, donc l'espérance de X est la somme de la famille $(x\mathbb{P}(X = x))_{x \in X(\Omega)}$. Il y a alors deux possibilités :

— ou bien la famille est sommable, et alors l'espérance et bien définie et appartient à $\mathbb{R}_+$.
— ou bien elle n'est pas sommable, et on pose alors par convention $E(X) = +\infty$.

Qualitativement : l'espérance doit être vue comme la valeur moyenne des valeurs de X (pondérées par leur probabilité d'apparition).

Pour définir l'espérance dans le cas général, on fait comme pour les familles sommables :

Définition (cas des variables aléatoires réelles discrètes)

Soit $X : (\Omega, \mathcal{A}, \mathbb{P}) \longrightarrow \mathbb{R}$ une variable aléatoire réelle discrète.
On dit que X possède une espérance si la famille $(x\mathbb{P}(X = x))_{x \in X(\Omega)}$ est sommable c'est à dire si

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

ou encore si $E(|X|) < +\infty$

Si c'est le cas on pose

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

.

Comme dans le cas des familles sommables, il faut faire attention à la place des valeurs absolues dans cette définition.

Exemple.

— Espérance des constantes.

— Espérance des indicatrices. Si A est un événement, on a

$$E(\mathbb{1}_A) = \mathbb{P}(A)$$

— Espérance des VA presque surement nulles.

— Espérance lors d'un lancer de dé.

Variables identiquement distribuées :
L'espérance d'une variable ne dépend que de sa loi. Si $X \sim Y$ on a donc $E(X) = E(Y)$.

Un exemple

Temps d'attente du premier succès dans un pile ou face. On tire à pile ou face (avec probabilité p pour obtenir pile). On note X la variable aléatoire égale au premier instant ou l'on obtient pile. Montrer que X est presque surement finie, d'espérance finie, et calculer son espérance.

II.2 Le théorème de transfert

Théorème

Soit $X : (\Omega, \mathcal{A}, \mathbb{P}) \longrightarrow E$ une variable aléatoire discrète. Soit $f : E \longrightarrow \mathbb{R}$. Alors la variable aléatoire réelle discrète $f(X)$ possède une espérance si et seulement si $(f(x)\mathbb{P}(X = x))_{x \in \mathbb{X}(\Omega)}$ est sommable. Dans ce cas :

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

Démonstration.

II.3 Linéarité de l'espérance

Théorème

Soient X et Y deux variables aléatoires discrètes admettant chacune une espérance. Alors $X+Y$ admet également une espérance, et $E(X+Y) = E(X) + E(Y)$.

Démonstration. Soient $X : \Omega \longrightarrow \mathbb{R}$ et $Y : \Omega \longrightarrow \mathbb{R}$. Posons $f : (u, v) \in \mathbb{R}^2 \longmapsto u + v$. Alors, en appliquant le théorème de transfert, $f(X, Y)$ possède une espérance si et seulement si $((x + y)\mathbb{P}((X, Y) = (x, y)))_{(x,y) \in X(\Omega) \times Y(\Omega)}$ est sommable. On note $\alpha_{x,y} = |x + y|\mathbb{P}((X, Y) = (x, y))$.

Remarquons que $\alpha_{x,y} \leqslant (|x|+|y|)\mathbb{P}(X=x, Y=y)$. Ainsi :

$$\begin{aligned}
&\sum_{\substack{x\in X(\Omega)\\ y\in Y(\Omega)}} |x+y|\mathbb{P}(X=x,Y=y)\\
&\leqslant \sum_{x\in X(\Omega)} |x| \sum_{y\in Y(\Omega)} \mathbb{P}(X=x,Y=y) + \sum_{y\in Y(\Omega)} |y| \sum_{x\in X(\Omega)} \mathbb{P}(X=x,Y=y)\\
&= \sum_{x\in X(\Omega)} |x|\mathbb{P}(X=x) + \sum_{y\in Y(\Omega)} |y|\mathbb{P}(Y=y)\\
&= E(|X|)+E(|Y|)
\end{aligned}$$

La famille est donc bien sommable, $X+Y$ possède une espérance, et en reprenant le même calcul (sans les valeurs absolues), qui est validé par la sommabilité, on montre que $E(X+Y)=E(X)+E(Y)$.

Exemple

Temps d'attente du deuxième succès dans un pile ou face. On tire à pile ou face (avec probabilité p pour obtenir pile). On note Z la variable aléatoire égale au second instant ou l'on obtient pile. Déterminer l'espérance de cette variable aléatoire , soit directement, soit en utilisant la linéarité de l'espérance.

Corollaire

Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé. L'ensemble $\mathcal{L}^1(\Omega, \mathcal{A}, \mathbb{P})$ des variables aléatoires réelles discrètes est un $\mathbb{R}$-espace vectoriel. De plus, l'application espérance est une forme linéaire sur cet espace.

De plus, la formule de transfert nous donne que $X \in \mathcal{L}^1(\Omega, \mathcal{A}, \mathbb{P})$ si et seulement si $|X| \in \mathcal{L}^1(\Omega, \mathcal{A}, \mathbb{P})$.

Proposition

i. L'espérance est une forme linéaire positive : si X est positive, alors $E(X) \geqslant 0$.
ii. $|E(X)| \leqslant E(|X|)$ pour toute VA ayant une espérance.
iii. Si X est positive d'espérance nulle, alors X est presque sûrement nulle.

Démonstration.

Théorème de comparaison

Soit X et Y deux variables aléatoires réelles discrètes, avec Y positive. On suppose que $|X| \leqslant Y$ et que Y admet une espérance. Alors X admet une espérance, et $E(|X|) \leqslant E(Y)$.

Proposition (variables bornées)

Toutes les variables aléatoires bornées discrètes possèdent une espérance, et si $|X|$ est majorée par M, alors $|E(X)| \leqslant M$.

‘

II.4 Moments d'ordre p

Définition

Soit X une variable aléatoire réelle discrète. On dit que X admet un moment d'ordre p (pour $p \in \mathbb{N}^*$) lorsque X^p possède une espérance. Alors, $E(X^p)$ s'appelle le moment d'ordre p de X.

Proposition

Si X possède un moment d'ordre p, alors pour tout $q \in [\![1, p]\!]$, $E(|X|^q)$ existe.

Démonstration.

II.5 Moments d'ordre 2, l'espace $\mathcal{L}^2$, Cauchy-Schwarz

L'espace $\mathcal{L}^2(\Omega, \mathcal{A}, \mathbb{P})$

Définition

On pose $\mathcal{L}^2(\Omega, \mathcal{A}, \mathbb{P}) = \{X,\ E(|X|^2) < +\infty\}$. Cet ensemble est un espace vectoriel.

Démonstration. Remarquons simplement que $|X+Y|^2 \leqslant 2(|X|^2 + |Y|^2)$.

Inégalité de Cauchy-Schwarz

Proposition

Considérons l'application ϕ suivante :

$$\phi : \left| \begin{array}{lcl} \mathcal{L}^2(\Omega, \mathcal{A}, \mathbb{P})^2 & \longrightarrow & \mathbb{R} \\ (X, Y) & \longmapsto & E(XY) \end{array} \right.$$

ϕ est une forme bilinéaire symétrique positive.

Théorème (Cauchy-Schwarz)

Soient X et Y des variables aléatoires réelles discrètes ayant un moment d'ordre 2. Alors

$$E(XY)^2 \leqslant E(X^2)E(Y^2)$$

Démonstration. Identique au cas des produits scalaires.

Exemples d'utilisation de l'inégalité de Cauchy Schwarz

L'inégalité de Cauchy Schwarz est encore plus importante pour les variables aléatoires que pour les fonctions. Il faut y penser dès que le carré des variables aléatoires intervient. En voici des exemples.

- Montrer que pour toute VA ayant un moment d'ordre 2 on a $E(X^2) \leqslant E(X)^2$.

- (plus subtil) Soit X une vard ayant un moment d'ordre 2 et X^+ sa partie positive. Montrer que X^+ a une espérance et que :

$$E(X^+) \leqslant \sqrt{\mathbb{P}(X \geqslant 0)E(X^2)}$$

II.6 Variance

Définition

Soit X une variable aléatoire réelle discrète admettant un moment d'ordre 2. On pose :

$$\begin{aligned} V(X) &= E\left((X - E(X))^2\right) \\ &= E(X^2) - E(X)^2 \quad \text{(Formule de König-Huygens)} \end{aligned}$$

La variance est un indicateur de dispersion par rapport à la moyenne comme l'explique la proposition ci-dessous.

Proposition

— Si a est une constante, $V(X + a) = V(X)$.
— $V(X) = 0 \iff X$ presque sûrement constante.

De plus on a la propriété algébriques suivante :

Proposition

La variance est homogène de degré 2 :

$$V(aX) = a^2V(X)$$

C'est la raison pour laquelle, sa racine carrée est intéressante :

Écart-type

Définition

Soit X une vard ayant une variance. On appelle l'écart-type de X la quantité $\sigma(X) = \sqrt{V(X)}$.

Cet indicateur est essentiel en statistiques (on l'appelle aussi l'écart quadratique moyen)

Variables aléatoires centrées réduites

On peut toujours renormaliser une variable aléatoire en faisant des transformation affines (translation pour changer la moyenne, homothétie pour changer l'écart type) :

Définition

— On dit que $X \in \mathcal{L}^1(\Omega, \mathcal{A}, \mathbb{P})$ est centrée lorsque $E(X) = 0$.
— On dit que $X \in \mathcal{L}^2(\Omega, \mathcal{A}, \mathbb{P})$ est réduite lorsque $V(X) = 1$.

Proposition

Soit $X^{\in}\mathcal{L}^2(\Omega, \mathcal{A}, \mathbb{P})$ une vard non presque sûrement constante. On pose :

$$Y = \frac{X - E(X)}{\sigma(X)}$$

Alors Y est centrée réduite.

II.7 Covariance

Définition

Soient X et Y deux variables aléatoires réelles discrètes ayant un moment d'ordre 2. On pose :

$$\operatorname{cov}(X, Y) = E((X - E(X))(Y - E(Y)))$$

Proposition

La covariance est une forme bilinéaire symétrique positive sur l'ensemble des variables admettant un moment d'ordre 2 de $(\Omega, \mathcal{A}, \mathbb{P})$. Elle vérifie donc Cauchy Schwarz.

Définition

Soit X et Y deux variables admettant un moment d'ordre 2. On définit $\rho(X, Y)$, le coefficient de corrélation de (X, Y) :

$$\rho(X, Y) = \frac{\operatorname{cov}(X, Y)}{\sigma(X)\sigma(Y)}$$

Par Cauchy-Schwarz, le coefficient de corrélation est inférieur à 1 en module. Cet indicateur mesure si les variables sont corrélées : plus il est proche de 1 (ou -1). plus les variables sont corrélées positivement (ou négativement). Lorsqu'il est nul on dit que les VA X et Y sont **non corrélées**.

III Propriété spécifique des vard à valeurs entières

III.1 Une formule pour l'espérance

Proposition

Soit X une vard à valeurs dans $\mathbb{N}$. Alors :

$$E(X) = \sum_{n=1}^{+\infty} \mathbb{P}(X \geqslant n)$$

Exemple : Temps d'attente du premier succès et du deuxième succès dans un pile ou face. On tire à pile ou face (avec probabilité p pour obtenir pile). On note X la variable aléatoire égale au premier instant ou l'on obtient pile. Déterminer l'espérance de cette variable aléatoire en utilisant la formule ci-dessus. (à comparer avec la méthode déjà employée).

III.2 La fonction génératrice

Définition

Soit X une vard à valeurs dans $\mathbb{N}$. Alors la série entière $\sum_{n\geqslant 0} \mathbb{P}(X = n)t^n$ converge normalement sur $[0, 1]$ (elle a donc a un rayon de convergence au moins égal à 1). La fonction :

$$G_X : t \in [0, 1] \longmapsto \sum_{n=0}^{+\infty} \mathbb{P}(X = n)t^n$$

s'appelle fonction (ou série) génératrice de X.

Démonstration.

L'intérêt de cette série entière est qu'elle **contient toutes les informations** sur la loi de la vard X.

Exemple : le lancer de dé
Soit X la variable aléatoire égale au résultat d'un lancer de dé à 6 faces. La fonction génératrice de X est :

III.3 Propriétés remarquables

Proposition

-Pour tout $t \in [0,1]$, $G_X(t) = E(t^X)$.

- G_X est continue sur $[0,1]$ fermé, en particulier au point 1.

-Si $G_X = G_Y$, alors X et Y suivent la même loi.

Démonstration.

III.4 Application à l'espérance et à la variance

Proposition

Il est équivalent de dire :

i. $E(X) < +\infty$

ii. G_X est dérivable en 1

Dans ce cas, $G'_X(1) = E(X)$.(cette formule doit être connue)

Il est équivalent de dire :

i. $V(X) < +\infty$

ii. G_X est deux fois dérivable en 1

Dans ce cas, $G''_X(1) = E(X(X-1))$.(cette formule permet de retrouver la variance sans difficulté)

Démonstration.

Proposition (Cas général)

X possède un moment d'ordre p si et seulement si G_X est p fois dérivable en 1.

La preuve se fait de la même façon.

IV Indépendance des variables aléatoires

IV.1 Définitions

Cas de deux variables aléatoires discrètes

Définition

Soient $X : (\Omega, \mathcal{A}, \mathbb{P}) \longrightarrow E$ et $Y : (\Omega, \mathcal{A}, \mathbb{P}) \longrightarrow F$. On dit que X et Y sont indépendants si :

$$\forall\, x \in X(\Omega),\ \forall\, y \in Y(\Omega),\ [X = x] \text{ et } [Y = y] \text{ sont indépendants.}$$

Cas général

Définition

Soit $(X_i)_{i\in I}$ une famille de vad. On dit qu'elles sont indépendantes lorsque $\forall\, n \in \mathbb{N}^*$, $\forall\, (i_1, ..., i_n) \in I^n$, $\forall\, (a_{i_1}, ..., a_{i_n})$, les événements $[X_{i_1} = a_{i_1}], ..., [X_{i_n} = a_{i_n}]$ sont mutuellement indépendants.

Dans les expériences telles que les tirages de pile ou face ou de dés, les variables aléatoires associées aux différents tirages seront systématiquement supposées indépendantes.
Exemple de référence :
Dans une suite de pile ou face, on note A_n la variable aléatoire égale au numero du tirage ou l'on obtient le $n$ième pile. Montrer que les variables aléatoires A_n sont indépendantes.

IV.2 Espérance et variance

En plus des propriétés usuelles de l'espérance et de la variance on a des résultats remarquables en cas d'indépendance. Les voici :

Proposition

Si X et Y sont indépendantes et ont une espérance :

$$E(XY) = E(X)E(Y)$$

Démonstration.

Corollaire

SI X et Y sont indépendantes et ont une variance alors :

a) $V(X+Y) = V(X) + V(Y)$

b) X et Y sont non corrélées : $\text{cov}(X,Y) = \rho(X,Y) = 0$ (réciproque fausse)

IV.3 Indépendance et séries génératrices

Proposition

Soient X et Y des vard à valeurs dans $\mathbb{N}$. On suppose que X et Y sont indépendantes. Alors

$$G_{X+Y} = G_X \cdot G_Y$$

IV.4 Lemme des coalitions

Le résultat suivant est très utile pour décider si des variables sont indépendantes.

Fonctions de variables indépendantes

Théorème

Soient X et Y deux vad indépendantes. Soient f et g deux fonctions définies sur $X(\Omega)$ et $Y(\Omega)$ réciproquement. Alors $f(X)$ et $g(Y)$ sont indépendantes.

et plus généralement :

Le lemme des coalitions

Théorème

Soient $(X_1, ..., X_n)$ une famille finie de variables aléatoires indépendantes. Soit $p \in [\![1, n]\!]$. Soient Y et Z des variables aléatoires de la forme $Y = f(X_1, ..., X_p)$ et $Z = g(X_{p+1}, ..., X_n)$. Alors Y et Z sont indépendantes.

Démonstration.

V Lois usuelles

V.1 La loi uniforme

Définition

On dit qu'une variable aléatoire X suit la loi uniforme $U(n)$ lorsque :

i. $X(\Omega) = [\![1, n]\!]$

ii. $\forall\, k \in [\![1, n]\!]$, $\mathbb{P}(X = k) = 1/n$

Proposition

— $G_X(t) = \frac{1}{n}(t + \cdots + t^n)$
— $E(X) = \frac{n+1}{2}$
— $V(X) = \frac{n^2-1}{12}$

l'exemple de référence est le lancer de dé.

Remarque. On peut également considérer la loi uniforme $U([\![a, b]\!])$, en opérant une translation de $a - 1$ (qui va affecter la moyenne mais pas la variance).

Notation : à la place de l'expression "X suit la loi uniforme $U(n)$" on écrit de façon simplifiée $X \hookrightarrow U(n)$. On fera de même pour toutes les lois usuelles.

V.2 La loi de Bernoulli

Définition

Soit $p \in [0, 1]$ et $q = 1 - p$. On dit que $X \hookrightarrow B(p)$ lorsque :

i. $X(\Omega) = \{0, 1\}$

ii. $\mathbb{P}(X = 1) = p$

Proposition

— $G_X(t) = pt + q$
— $E(X) = p$
— $V(X) = pq$

La loi de Bernoulli modélise les tirages de pile ou face, ou plus généralement, tous les tirages ayant deux issues. Elle est très importante.

Variante de la loi de Bernoulli . La loi de **Rademacher** est donnée par :

i. $X(\Omega) = \{-1, 1\}$

ii. $\mathbb{P}(X = 1) = \dfrac{1}{2}$

Cette loi sert à modéliser les marches aléatoires symétriques (on diminue ou augmente de 1 avec équiprobabilité). Cette loi n'a pas de série génératrice car elle n'est pas à valeur dans $\mathbb{N}$, mais elle a espérance et variance bien entendu.

V.3 La loi binomiale $B(n,p)$

Proposition

Soit $(X_1, ..., X_n)$ un $n-$uplet de variables indépendantes suivant la loi $B(p)$.
Alors la variable aléatoire $X = X_1 + \cdots + X_n$ vérifie :

i. $X(\Omega) = [\![0, n]\!]$

ii. $\mathbb{P}(X = k) = \binom{n}{k} p^k q^{n-k}$ pour $k \in [\![1, n]\!]$ avec $q = 1 - p$

On dit que la variable X suit une loi binomiale de paramètres (n,p) :

$$X \hookrightarrow B(n,p)$$

Démonstration. On conclut par théorème d'unicité.

La loi $B(n,p)$ modélise le nombre de succès dans la répétition de n expériences de Bernoulli indépendantes, de paramètre p.

Proposition

— $G_X(t) = (pt + q)^n$
— $E(X) = np$
— $V(X) = npq$

Additivité des variables binomiales indépendantes

Proposition

Soit (X,Y) un couple de variables indépendantes avec $X \hookrightarrow B(n,p)$ et $Y \hookrightarrow B(m,p)$. Alors :

$$X + Y \hookrightarrow B(m+n, p)$$

V.4 La loi géométrique

Définition

On dit qu'une variable aléatoire X définie sur $(\Omega, \mathcal{A}, \mathbb{P})$ suit $\mathcal{G}(p)$, la loi géométrique de paramètre p lorsque :

i. $X(\Omega) = \mathbb{N}^*$
ii. $\forall\, k \in \mathbb{N}^*,\ \mathbb{P}(X = k) = pq^{k-1}$

Elle modélise le premier succès dans la répétition d'expériences de Bernoulli indépendantes de paramètre p.

Proposition

— $G_X(t) = \frac{pt}{1-qt}$
— $E(X) = \frac{1}{p}$
— $V(X) = \frac{q}{p^2}$

Proposition (loi sans mémoire)

La loi géométrique est caractérisée par la propriété suivante (on dit qu'elle est sans mémoire)
Pour tout n, pour tout k,
$$\mathbb{P}(X > n + k | X > n) = \mathbb{P}(X > k)$$
Réciproquement, si X est non presque sûrement constante avec $X(\Omega) \subset \mathbb{N}^*$, alors il existe p tel que $X \hookrightarrow \mathcal{G}(p)$.

V.5 La loi de Poisson

Définition

Soit $\lambda \in \mathbb{R}_+^*$, on dit que la variable aléatoire X suit une loi de Poisson de paramètre λ ($X \hookrightarrow \mathcal{P}(\lambda)$) lorsque :

i. $X(\Omega) = \mathbb{N}$
ii. $\forall\, n \in \mathbb{N},\ \mathbb{P}(X = n) = \frac{\lambda^n}{n!} e^{-\lambda}$.

Les lois de Poisson modélisent généralement les files d'attentes, ou le nombre de succès dans un laps de temps fixe.

Proposition

- $G_X(t) = e^{\lambda(t-1)}$
- $E(X) = V(X) = \lambda$

Additivité des variables de Poisson

Proposition

Soit (X, Y) un couple de variables indépendantes avec $X \hookrightarrow \mathcal{P}(\lambda)$ et $Y \hookrightarrow \mathcal{P}(\mu)$. Alors :

$$X + Y \hookrightarrow \mathcal{P}(\lambda + \mu)$$

Proposition (Approximation de la loi binomiale pour les événements rares)

Soit (X_n) une suite de variables aléatoires suivant la loi $B(n, p_n)$. Supposons que $np_n \xrightarrow[n\to+\infty]{} \lambda$.
Alors $\mathbb{P}(X_n = k) \xrightarrow[n\to+\infty]{} \mathbb{P}(X = k)$ avec $X \hookrightarrow \mathcal{P}(\lambda)$.

La loi de Poisson permet de faire une approximation de la loi binomiale pour les grandes valeurs de n et les petites valeurs de p avec le paramètre $\lambda = np$
Par exemple : On tire 520 fois une carte avec remise dans un jeu de 52. Quelle est la probabilité d'obtenir 10 fois l'as de pique ?

Ici $n = 520; p = 1/52$ la probabilité cherchée est $\frac{(51)^{510}}{52^{520}}\binom{520}{10}$ qui vaut environ 0.126330 l'approximation de Poisson est $\frac{10^{10}}{10!}e^{-10}$. La formule de Stirling approxime ceci en $\frac{1}{\sqrt{20\pi}}$ qui vaut environ 0.126156

VI Vecteurs aléatoires, lois conjointes, lois marginales

VI.1 Loi d'un couple

Définition

Soient X, Y deux variables aléatoires discrètes définies sur $(\Omega, \mathcal{A}, \mathbb{P})$. On appelle loi conjointe des variables aléatoires la loi du couple (X, Y) donnée par $(\mathbb{P}(X = x, Y = y))_{x \in X(\Omega), y \in Y(\Omega)}$.
On appelle lois marginales les lois de X et de Y.

On peut ceci sous forme d'un tableau :

	x_1	$\dots$	x_i	$\dots$	x_n	$X(\Omega)$
y_1						
$\vdots$						
y_j			$P(X = x_i, Y = y_j)$			$P(Y = y_j)$
$\vdots$						
y_p						
$Y(\Omega)$			$P(X = x_i)$			1

En sommant sur les lignes on trouve la loi de Y, en sommant sur les colonnes on trouve la loi de X. La somme de tous les éléments du tableau vaut 1.

Proposition

La loi conjointe détermine les lois marginales, mais la réciproque est fausse.

VI.2 Caractérisation de l'indépendance

Proposition

Si X et Y sont indépendantes, les lois marginales déterminent les lois conjointes

Plus précisément le tableau ci dessus est entièrement déterminé par la dernière colonne et la dernière ligne, puisque $\mathbb{P}(X = x, Y = y) = \mathbb{P}(X = x)\mathbb{P}(Y = y)$.

Ceci est faux si les variables ne sont pas indépendantes.

Cette définition s'étend sans difficulté :

Définition (cas de n variables)

Soit $(X_1, ..., X_n)$ un n-uplet de variables aléatoires discrètes. La loi conjointe est donnée par $\mathbb{P}(X_1 = x_1, ..., X_n = x_n)$ pour $(x_1, ..., x_n) \in X_1(\Omega) \times \cdots \times X_n(\Omega)$.

Dans ce cas aussi, la loi conjointe est déterminée par les lois marginales dans le cas des variables indépendantes, mais pas en général.

VI.3 Loi conditionnelle

Définition

Soit $A \in \mathcal{A}$. On appelle loi conditionnelle de X relativement à l'événement A la loi de probabilité $\mathbb{P}_X(\cdot|A)$ définie par :

$$\mathbb{P}_X(B|A) = \mathbb{P}(X \in B|A)$$

VI.4 Etude complète d'un exemple

Variante autour de l'exercice 98 du poly d'oral CCP.
Soit n un entier. On effectue n épreuves de Bernoulli de paramètre p. Soit X le nombre de succès. On effectue ensuite $n - X$ épreuves de Bernoulli. On note Y le nombre de nouveaux succès.
Déterminer :
a) La loi de X
b) La loi de Y conditionnellement à l'événement $X = i$
c) La loi conjointe du couple (X, Y).
d) si les variables X, Y sont indépendantes ou non.
e) La loi de Y.

VII Inégalité de Markov, de Bienaymé-Tchebychev, loi faible des grands nombres

VII.1 Inégalité de Markov

Proposition

Soit X une variable aléatoire à valeurs positives et ayant une espérance. Alors :

$$\forall\, a \in \mathbb{R}_+,\ \mathbb{P}(X \geqslant a) \leqslant \frac{E(X)}{a}$$

Démonstration.

Remarque. On utilise souvent une amélioration de l'inégalité de Markov :

Proposition

Si f est une fonction strictement croissante et strictement positive telle que $f(X)$ est d'espèrance finie alors :

$$\mathbb{P}(X \geqslant a) \leqslant \frac{E(f(X))}{f(a)}$$

VII.2 Inégalité de Bienaymé-Tchebychev

Proposition

Soit X une variable aléatoire ayant un moment d'ordre 2. Soit $\epsilon > 0$.

$$\mathbb{P}(|X - E(X)| \geqslant \epsilon) \leqslant \frac{V(X)}{\epsilon^2}$$

Démonstration.

Ces inégalités sont utilisées pour contrôler l'écart entre une variable aléatoire et sa moyenne. Nous en verrons des utilisations en exercice.
Par exemple : on tire 1000 fois à pile ou face (donc en moyenne 500 pile). Majorer la probabilité de faire au moins 550 pile par l'inégalité de Bienaymé Tchebychev.

L'application théorique la plus importante de ces inégalités est le résultat ci-dessous.

VII.3 Loi faible des grands nombres

On se donne une variable aléatoire X qui modélise une certaine expérience. On suppose l'existence d'une suite $(X_n)_{n\in\mathbb{N}}$ de variables aléatoires définies sur un espace probabilisé telle que :
— pour tout $n \in \mathbb{N}$, $X_n \sim X$
— les X_n sont mutuellement indépendantes
On dit alors que (X_n) est une suite de variables aléatoires indépendantes identiquement distribuées (iid).

Théorème (Loi faibles des grands nombres)

Soit (X_n) une suite de variables aléatoires discrètes indépendantes identiquement distribuées suivant la loi de X. On suppose que X possède une variance. Alors :

$$\forall\, \epsilon > 0,\ \mathbb{P}\left(\left|\frac{X_1 + \cdots + X_n}{n} - E(X)\right| \geqslant \epsilon\right) \xrightarrow[n\to+\infty]{} 0$$

Démonstration.

VII.4 Types de convergences(HP)

Soit (X_n) une suite de variables aléatoires réelles discrètes. On distingue plusieurs type de convergences de (X_n) vers une éventuelle variable aléatoire limite X :
— La convergence en loi :

$$\forall\, (a,b) \in \mathbb{R}^2,\ \mathbb{P}(X_n \in [a,b]) \xrightarrow[n\to+\infty]{} \mathbb{P}(X \in [a,b])$$

— La convergence en probabilité :

$$\forall\, \epsilon > 0,\ \mathbb{P}(|X_n - X| \geqslant \epsilon) \xrightarrow[n\to+\infty]{} 0$$

— La convergence presque sûre :

$$\mathbb{P}\left(X_n \xrightarrow[n\to+\infty]{} X\right) = 1$$

L'étude de ces différentes notion n'est pas au programme. Avec ce vocabulaire, la loi faible des grands nombres exprime que la suite de variables aléatoires $\frac{X_1+\cdots+X_n}{n}$ converge en probabilité vers sa moyenne. C'est un fait très intuitif.

A titre d'exercice, comment s'interprète l'approximation d'une loi binomiale par une loi de Poisson en termes de convergences de variables aléatoires ?