Ecole Nationale des Travaux Publics de l'Etat
3 Rue Maurice Audin
69 518, Vaulx en Velin cedex, France

Probabilités

Notes de cours et Exercices

Première année

Chapitre 1

Modélisation des phénomènes aléatoires. Espaces probabilisés

Chapitre 2

Conditionnement et Indépendance

Chapitre 3

Variables et vecteurs aléatoires discrets

1 Introduction

Après avoir réalisé une expérience, il arrive bien souvent qu'on s'intéresse plus à une fonction du résultat qu'au résultat en lui-même. Par exemple, lors d'un lancer de deux dés, on regarde la somme des dés ou bien lors de la répétition du lancer d'un epièce, on regarde le nombre de fois où l'on obtient "pile".
Cette fonction du résultat est souvent à valeurs dans $\mathbb{R}$ ou dans $\mathbb{R}^n$ et on parle alors de variable respectivement de vecteur aléatoire.

2 Généralités

Soit $(\Omega, \mathscr{T}, P)$ un espace probabilisé.

Rappel : Soit X une application de $\Omega \to \mathbb{R}$. Pour tout $B \subset \mathbb{R}$, on note :

$$\{X \in B\} := \{w \in \Omega : X(w) \in B\} := X^{-1}(B)$$

l'ensemble des antécédents des éléments de B par l'application X.
De plus, si $(B_i)_{i \in I}$ est une famille quelconque de parties de $\mathbb{R}$, on a :

$$X^{-1}\left(\bigcup_{i \in I} B_i\right) = \bigcup_{i \in I} X^{-1}(B_i),\ X^{-1}\left(\bigcap_{i \in I} B_i\right) = \bigcap_{i \in I} X^{-1}(B_i).$$

Enfin, pour tout $B \subset \mathbb{R}, X^{-1}(\mathbb{R} \setminus B) = \Omega \setminus X^{-1}(B)$.

2.1 Variable aléatoire

Une **variable aléatoire** définie sur Ω est une appplication $X : \Omega \to \mathbb{R}$ telle que, pour tout borélien B de $\mathbb{R}$, l'ensemble $X^{-1}(B) = \{w \in \Omega : X(w) \in B\}$ est un évènement, i.e. $X^{-1}(B) \in \mathscr{T}$.

On dit alors que X est mesurable pour la tribu des boréliens de $\mathbb{R}$ et la tribu $\mathscr{T}$. On peut montrer que pour que $X : \Omega \to \mathbb{R}$ soit une variable aléatoire, il suffit que :

$$\forall a, b \in \mathbb{R}, \quad X^{-1}(]a, b]) \in \mathscr{T}.$$

Remarque :
- Une variable aléatoire, malgré son nom, n'est pas une variable (au sens de l'analyse), mais une fonction de la variable $w \in \Omega$.
- lorsque la tribu $\mathscr{T}$ est $\mathcal{P}(\Omega)$, comme dans le cas d'un ensemble Ω fini ou dénombrable, toute application $X : \Omega \to \mathbb{R}$ est une variable aléatoire.

2.2 Vecteur aléatoire

Un **vecteur aléatoire** définie sur Ω est une appplication $X : \Omega \to \mathbb{R}^n$ telle que, pour tout borélien B de $\mathbb{R}^n$, l'ensemble $X^{-1}(B) = \{w \in \Omega : X(w) \in B\}$ est un évènement, i.e. $X^{-1}(B) \in \mathscr{T}$.

Pour celà, on montre qu'il suffit en fait que $\{X \in J\} \in \mathscr{T}$ pour tout ensemble J de $\mathbb{R}^n$ de la forme $J = \prod_{k=1}^{n} I_k$ où, pour tout k, I_k est un intervalle de $\mathbb{R}$.

Remarque : Lorsque la tribu $\mathscr{T}$ est $\mathcal{P}(\Omega)$, toute application $X : \Omega \to \mathbb{R}^n$ est un vecteur aléatoire.

3 Variables et vecteurs aléatoires discrets

3.1 Définition

Une variable aléatoire $X : \Omega \to \mathbb{R}$ est dite **discrète** si l'ensemble $S = X(\Omega)$ est une partie finie ou dénombrable de $\mathbb{R}$.
Cette définition s'étend au cas d'un **vecteur aléatoire** qui est dit **discret** si l'ensemble $V = X(\Omega)$ des valeurs de X est une partie finie ou dénombrable de $\mathbb{R}^n$.

Remarque :
- Ainsi, $X = (X_1, X_2, ..., X_n) : \Omega \to \mathbb{R}^n$ est un vecteur aléatoire discret si et seulement si ses applications coordonnées $X_k : \Omega \to \mathbb{R}, 1 \leq k \leq n$ sont des variables aléatoires discrètes.
- Le couple $(X, Y) : \Omega \to \mathbb{R}^2$ est discret si et seulement si les deux variables aléatoires $X : \Omega \to \mathbb{R}$ et $Y : \Omega \to \mathbb{R}$ sont des variables aléatoires discrètes. De plus si $S = X(\Omega)$ et $T = Y(\Omega)$ alors $(X, Y)(\Omega) \subset S \times T$.

Exemples

1. On lance un dé équilibré : on gagne 1 euro si le numéro est pair, on perd 1 euro sinon. Ceci peut être modélisé par l'application $X : (\Omega, \mathcal{P}(\Omega), P) \to \{-1, +1\}$ avec $\Omega = \{1, 2, 3, 4, 5, 6\}$ et P l'équiprobabilité sur Ω. Les deux évènemens $\{X = +1\} = \{2, 4, 6\}$

et $\{X = -1\} = \{1, 3, 5\}$ appartiennent à $\mathcal{P}(\Omega)$, c'est-à-dire que X est bien une variable aléatoire discrète.

2. On lance deux dés équilibrés et on s'intéresse à leur somme. Celle-ci est une variable aléatoire discrète X à valeurs dans $\{2, ..., 12\}$,$\Omega = \{(i, j), 1 \leq i, j \leq 6\}$ étant muni de la tribu $\mathcal{P}(\Omega)$ et de l'équiprobabilité P.

3.2 Loi d'une variable aléatoire discrète

La **loi d'une variable aléatoire discrète** X (encore appelée distribution) à valeurs dans un ensemble fini ou dénombrable $S = X(\Omega)$ est caractérisée par $\{(s, p_s), s \in S\}$ avec

$$p_s = P_X(\{s\}) = P(w, X(w) = s) = P(X = s), \forall s \in S.$$

Bien entendu,

$$\forall s,\ 0 \leq p_s \leq 1 \text{ et } \sum_{s \in S} p_s = 1.$$

Autrement dit, la loi d'une variable aléatoire disctrète est complètement déterminée quand on précise l'ensemble des valeurs de la variable et la probabilité avec laquelle cette variable prend chacune de ses valeurs.

Exemples

1. Lancer d'un dé : la loi de X est donnée par $p_{-1} = p_{+1} = 1/2$. On dit que X suit une loi uniforme sur l'ensemble $\{-1, +1\}$.
2. Lancer de deux dés : la loi de X est donnée par le vecteur ligne $p = [p_2, ..., p_{12}]$ suivant :

$$p = \left[\frac{1}{36}, \frac{2}{36}, \frac{3}{36}, \frac{4}{36}, \frac{5}{36}, \frac{6}{36}, \frac{5}{36}, \frac{4}{36}, \frac{3}{36}, \frac{2}{36}, \frac{1}{36}\right]$$

3.3 Egalité en loi

Deux variables aléatoires discrètes X et Y sont dites identiquement distribuées, ou égales en loi et on note $X \sim Y$ ou $X \stackrel{\mathcal{L}}{=} Y$ si elles admettent la même loi, i.e. si $X(\Omega) = Y(\Omega) = S$ et si, pour tout $s \in S$, $P(X = s) = P(Y = s) = p_s$

Remarque : Deux variables alétaoires peuvent avoir même loi sans être égales. Par exemple, si on considère le jet de deux dés, l'un bleu et l'autrte rouge. On note X le nombre de points indiqué par le dé bleu et Y celui du rouge. Les variables aléatoires X et Y sont définies sur le même espace probabilisé $\Omega = \{1, 2, 3, 4, 5, 6\}^2$ muni de l'équiprobabilité.
On a $X(\Omega) = Y(\Omega) = \{1, 2, 3, 4, 5, 6\}$ et :

$$\forall k \in \{1, 2, 3, 4, 5, 6\},\ P(X = k) = \frac{1}{6},\ P(Y = k) = \frac{1}{6}.$$

Donc X et Y ont même loi. Pour autant on n'a pas l'égalité des variables aléatoires X et Y qui signifierait $X(w) = Y(w)$ pour tout $w \in \Omega$. Autrement dit, en lançant deux dés, on obtiendrait à coup sûr un double.

3.4 Loi d'un couple de variables aléatoires discrètes

Soit (X, Y) un couple de variables aléatoires discrètes défini sur un espace probabilisé $(\Omega, \mathscr{T}, P)$, $S = X(\Omega)$, $T = Y(\Omega)$ et $V = S \times T$. On considère que l'ensemble fini ou dénombrable V est l'ensemble des valeurs de (X, Y) de sorte que la loi du couple (encore appelée distribution) est caractérisée par $\{(v, p_v), v \in V\}$ avec

$$p_v = p_{s,t} = P((X, Y) = v) = P(X = s, Y = t), \ \forall v \in V = S \times T.$$

Remarque : La loi du couple (X, Y) est encore appelée **loi conjointe ou distribution conjointe**.

Disposition pratique : Lorsque S et T sont finis, la loi conjointe est présentée dans un tableau à double entrée où figurent les nombres $p_{s,t}$. L'image et la loi de X occupent la première et la dernière colonne, celles de Y occupent la première et la dernière ligne.

Exemple On répartit au hasard 3 boules numérotées de 1 à 3 dans 3 tiroirs numérotés de 1 à 3. On note X le nombre de boules dans le tiroir 1 et Y le nombre de tiroirs vides. On veut déterminer la loi conjointe et les lois marginales de (X, Y).
On modélise cette expérience en prenant pour Ω l'ensemble des répartitions des boules dans les tiroirs donc $\Omega = \{1, 2, 3\}^3$ et on considère la probabilité uniforme sur Ω. On a donc $S = X(\Omega) = \{0, 1, 2, 3\}$ et $T = \{0, 1, 2, \}$ et les probabilités $p_{i,j} = P(X = i, Y = j)$ se calculent facilement en divisant par 27 le nombre de cas favorables à l'évènement. On obtient ainsi le tableau suivant :

$X \backslash Y$	0	1	2	Loi de X
0	$0/27$	$6/27$	$2/27$	$8/27$
1	$6/27$	$6/27$	$0/27$	$12/27$
2	$0/27$	$6/27$	$0/27$	$6/27$
3	$0/27$	$0/27$	$1/27$	$1/27$
Loi de Y	$6/27$	$18/27$	$3/27$	1

3.5 Lois marginales

Les lois des coordonnées X et Y sont aussi appelées **lois marginales ou distributions marginales**. Ces lois se déduisent de la loi conjointe comme suit :

1. La loi de X est la loi sur S définie par les nombres

$$p_s = P(X = s) = p_{s.} = \sum_{t \in T} p_{s,t}$$

2. La loi de Y est la loi sur T définie par les nombres

$$q_t = P(Y = t) = p_{.t} = \sum_{s \in S} p_{s,t}$$

Remarque : La connaissance de la loi du couple permet de calculer les lois marginales. En revanche, la réciproque est fausse comme le montre l'exemple suivant.
On jette un dé bleu et un rouge. On note X les points indiqués par le dé bleu, Y ceux du dé rouge et on pose $Z = 7 - X$. Alors les couples aléatoires (X, Y) et (X, Z) ont mêmes lois marginales mais pas même loi. En effet il est clair que X, Y et Z suivent chacune la loi uniforme sur $\{1, 2, 3, 4, 5, 6\}$. Les lois des couples sont données par les tableaux suivants

	1	2	3	4	5	6
1	$1/36$	$1/36$	$1/36$	$1/36$	$1/36$	$1/36$
2	$1/36$	$1/36$	$1/36$	$1/36$	$1/36$	$1/36$
3	$1/36$	$1/36$	$1/36$	$1/36$	$1/36$	$1/36$
4	$1/36$	$1/36$	$1/36$	$1/36$	$1/36$	$1/36$
5	$1/36$	$1/36$	$1/36$	$1/36$	$1/36$	$1/36$
6	$1/36$	$1/36$	$1/36$	$1/36$	$1/36$	$1/36$

Loi de (X, Y)

	1	2	3	4	5	6
1	0	0	0	0	0	$1/6$
2	0	0	0	0	$1/6$	0
3	0	0	0	$1/6$	0	0
4	0	0	$1/6$	0	0	0
5	0	$1/6$	0	0	0	0
6	$1/6$	0	0	0	0	0

Loi de (X, Z)

3.6 Lois conditionelles

1. Soit $s \in S$ tel que $p_s = P(X = s) > 0$. La **loi conditionnelle** de Y sachant $X = s$ est la loi sur T définie par les nombres :

$$p_{t|s} = P(Y = t | X = s) = \frac{p_{s,t}}{p_{s.}}$$

et on a bien $\sum\limits_{t \in T} p_{t|s} = 1$

2. Soit $t \in T$ tel que $q_t = P(Y = t) > 0$. La **loi conditionnelle** de X sachant $Y = t$ est la loi sur S définie par les nombres :

$$p_{s|t} = P(X = s | Y = t) = \frac{p_{s,t}}{p_{.t}}$$

et on a bien $\sum\limits_{s \in S} p_{s|t} = 1$

Avec ces deux relations, on montre que

$$\forall (s,t) \in S \times T,\ p_{s,t} = p_{t|s} p_s = p_{s|t} q_t$$

Nous allons maintenant nous intéresser à une fonction qui permet de caractériser aussi bien les valeurs prises par une variable aléatoire discrète que sa loi : la fonction de répartition.

3.7 Fonction de répartition d'une variable aléatoire discrète

a) Définition

Soit X une variable aléatoire discrète et $S = X(\Omega) \subset \mathbb{R}$ l'ensemble de ses valeurs. La fonction de répartition de X est la fonction $F_X : \mathbb{R} \to [0,1]$ définie par :

$$F_X(x) = P(X \leq x) = \sum_{s \in S : s \leq x} p_s$$

b) Propriétés

Soit X une variable aléatoire discrète à valeurs dans S. Sa fonction de répartition F_X a les propriétés suivantes :

1. F_X est croissante ;
2. F_X est continue à droite ;
3. $\lim\limits_{x \to -\infty} F_X(x) = 0,\ \lim\limits_{x \to \infty} F_X(x) = 1$
4. $\forall x \in \mathbb{R},\ P(X = x) = F_X(x) - F_X(x^-)$ où $F_X(x^-) = \lim\limits_{\delta \to 0^+} F_X(x - \delta)$

Remarque :

– La dernière relation montre que la fonction de répartition d'une variable aléatoire discrète détermine complètement la loi de cette variable. Les seuls endroits où F_X présente des sauts sont ceux où X a des chances de tomber, la hauteur de chaque saut étant égale à la probabilité de tomber en ce point.

– A la place de la fonction de répartition F_X d'une variable aléatoire disctrète X, il est parfois utile de considérer la fonction G définie par

$$G_X(x) = 1 - F_X(x) = P(X > x)$$

qui est une fonction décroissante ayant des propriétés analogues à celles de F_X.

Exemples :

1. Lancer d'un dé : F_X ne prend que 3 valeurs, à savoir 0 sur $]-\infty, -1[$, $1/2$ sur $[-1, +1[$ et 1 sur $[1, +\infty[$.
2. Lancer de deux dés : F_X est une fonction en escalier (voir figure 3.1). Elle vaut 0 sur $]-\infty, 2[$, $1/36$ sur $[2, 3[$, $3/36$ sur $[3, 4[$, ..., $35/36$ sur $[11, 12[$ et 1 sur $[12, +\infty[$.

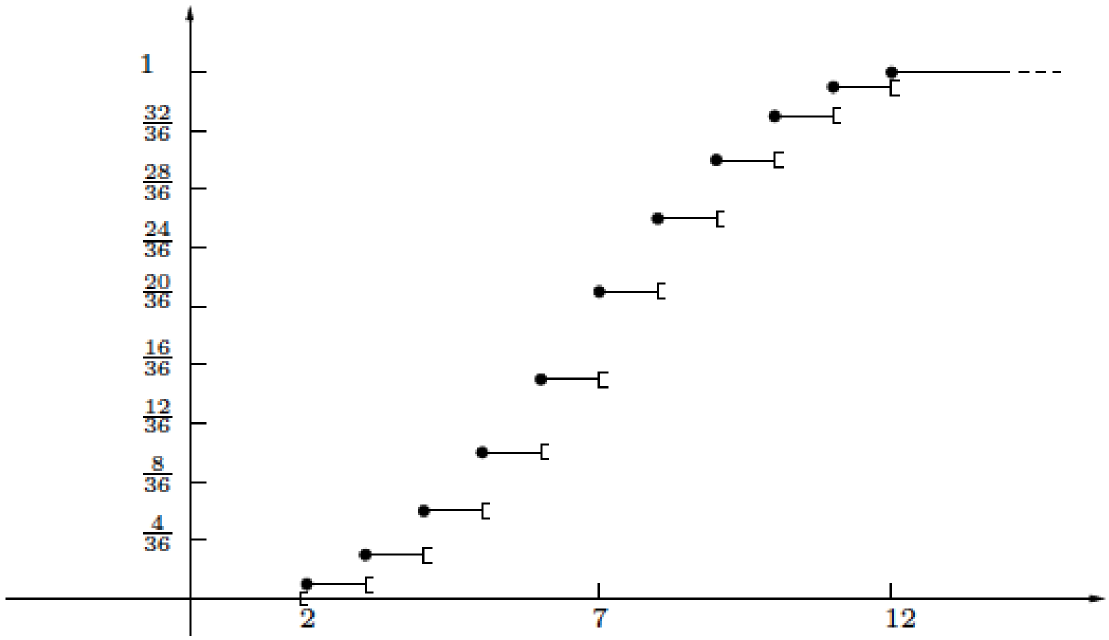

FIGURE 3.1 – Fonction de répartition pour la somme de deux dés.

3.8 Fonction de répartition d'un couple de variables aléatoires discrètes

Soit $(X, Y) : \Omega \to \mathbb{R}^2$ un couple de variables aléatoires discrètes, $S = X(\Omega)$, $T = Y(\Omega)$ et $V = S \times T$.
La fonction de répartition de (X, Y) est la fonction $F : \mathbb{R}^2 \to [0, 1]$ définie par

$$F(x, y) = P(X \leq x, Y \leq y) = \sum_{(s,t)\in V: s\leq x, t\leq y} p_{s,t}$$

4 Moments d'une variable aléatoire discrète

4.1 Espérance ou moment d'ordre 1

a) Définition
On dit que la variable X est d'ordre 1 lorsque la somme $\sum_{s\in S} |s|p_s$ est convergente.
Si tel est le cas, on appelle **espérance** de X et on note $\mathbb{E}(X)$ la quantité :

$$m = \mathbb{E}(X) = \sum_{s\in S} s\, p_s$$

L'ensemble des variables aléatoires discrètes d'ordre 1 est un espace vectoriel noté $L^1_d(\Omega)$.

Interprétation : L'espérance de X peut être vue comme la moyenne des valeurs s pondérées par les probabilités p_s, c'est pourquoi on dit aussi moyenne de X pour parler de son espérance.

Remarque :
- Lorsque S est fini, la moyenne est toujours bien définie.
- Si S est infini mais contenu dans $\mathbb{R}^+$ ou dans $\mathbb{R}^-$, la vérification se fait en même temps que le calcul de l'espérance.
- Soit $A \in \mathscr{T}$ et $\mathbb{1}_A$ sa fonction indicatrice. $\mathbb{1}_A$ est une variable aléatoire discrète qui prend la valeur 1 avec la probabilité $P(A)$ et la valeur 0 avec la probabilité $1 - P(A)$. En appliquant la définition, on voit que $\mathbb{E}(\mathbb{1}_A) = P(A)$.

Exemples :

1. Lancer d'un dé : le gain moyen est
$$\mathbb{E}(X) = (-1)\,P(X = -1) + 1\,P(X = 1) = 0$$
En moyenne, vous ne perdrez ni ne gagnerez donc rien à ce jeu (rien d'étonnant).
2. Un nombre est choisi au hasard entre 1 et 10, et on doit deviner ce nombre en posant des questions auxquelles il ne sera répondu que par oui ou par non. Calculons l'espérance du nombre N de questions nécessaires dans les cas suivants :
   - La ième question est du type "Est-ce i ?", i étant égal à $1, 2, ..., 10$:
   Soit A_k l'évènement : " le nombre $k \in \{1, ...10\}$ a été choisi". Alors
   $$P(N = k) = P(N = K|A_k)P(A_k) = 1/10$$
   et
   $$\mathbb{E}(N) = \sum_{k=1}^{10} kP(N = k) = 11/2$$
   - Avec chaque question, nous essayons d'éliminer à peu près la moitié des réponses possibles, avec le protocole suivant : est-ce ≤ 5, ≤ 2 (resp ≤ 7), ≤ 4 (resp ≤ 9) alors
   $$\mathbb{E}(N) = 3 \times 6/10 + 4 \times 4/10 = 17/5$$

b) Propriétés de la moyenne

Soient X et Y deux variables aléatoires discrètes d'ordre 1. On a les propriéts suivantes :

1. Pour toute constante c, on a $\mathbb{E}(c) = c$
2. $a \leq X \leq b \Rightarrow a \leq \mathbb{E}(X) \leq b$. En particulier, $X \geq 0 \Rightarrow \mathbb{E}(X) \geq 0$.
3. *Linéarité*. Pour tout couple (a, b) de scalaires, $aX + bY$ est d'orddre 1 et
$$\mathbb{E}(aX + bY) = a\mathbb{E}(X) + b\mathbb{E}(Y)$$
4. $X \leq Y \Rightarrow \mathbb{E}(X) \leq \mathbb{E}(Y)$

5. $|\mathbb{E}(X)| \leq \mathbb{E}(|X|)$

Exemple : min et max. On jette simultanément deux dés et on appelle X (resp. Y) le minimum (resp. le maximum) des deux numéros obtenus. Il est bien clair que $X \leq Y$ et le résultat ci-dessus assure, sans aucun calcul, que $\mathbb{E}(X) \leq \mathbb{E}(Y)$.

4.2 Moment d'ordre $k, k \geq 1$

a) Variable d'ordre k
Soit k un entier supérieur ou égal à 1. On dit que la variable X est d'ordre k lorsque la somme $\sum_{s\in S} |s|^k p_s$ est convergente.

b) Moment d'ordre k
Soit X une variable aléatoire discrète d'ordre k alors le moment d'ordre k de X est défini par :

$$m_k = \mathbb{E}(X^k) = \sum_{s\in S} s^k p_s$$

c) Moment centré d'ordre k
Soit X une variable aléatoire discrète d'ordre k alors le moment centré d'ordre k de X est défini par :

$$\mathbb{E}((X-m)^k) = \sum_{s\in S} (s-m)^k p_s$$

où m est l'espérance de X.

Remarque :
– Lorsque X est bornée ou lorsque S est fini alors X admet des moments de tout ordre.
– Si X est d'ordre k alors X est d'ordre j pour tout $j \leq k$.

d) Inégalité de Cauchy-Schwarz
Soient X et Y deux variables aléatoires discrètes d'ordre 2. Alors la variable aléatoire XY est d'ordre 1 et on a :

$$|\mathbb{E}(XY)| \leq \mathbb{E}(|XY|) \leq (\mathbb{E}(X^2)^{1/2}(\mathbb{E}(Y^2)^{1/2}$$

4.3 Variance et écart-type d'une variable aléatoire discrète

Considérons un amphi de 100 étudiants venant de subir un D.S. où la moyenne de l'amphi a été 10. On choisit un étudiant au hasard et on désigne par X sa note ; $X(\Omega)$ est l'ensemble des notes (toutes différentes) attribuées à ce D.S.. Comme tous les choix d'étudiants sont équiprobables, pour tout $x_k \in X(\Omega)$,

$$P(X = x_k) = \frac{\text{Nombre d'étudiants ayant obtenu la note} x_k}{100}.$$

D'où : $\mathbb{E}(X) = \sum_{x_k \in X(\Omega)} x_k P(X = x_k) = 10$. Cette moyenne ne nous apporte pas une information très précise sur l'amphi. Elle peut être obtenue avec 100 étudiants ayant 10 aussi bien qu'avec 50 ayant 0 et 50 ayant 20 ou un grand nombre de situations intermédiaires entre ces deux configurations extrêmes. Il est important pour compléter l'information apportée par la moyenne de disposer d'une quantité permettant de mesurer la dispersion autour de cette moyenne. C'est ce qu'on va définir maintenant.

Soit X une variable aléatoire discrète d'ordre 2 d'espérance m.

a) Variance

On appelle **variance** de X la quantité définie par

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

b) Ecart-type

On appelle **écart-type** de X la quantité définie par

$$\sigma = (\mathbb{V}(X))^{1/2}$$

c) Variable centrée réduite

Une variable aléatoire discrète Z est dite **variable centrée réduite** si $\mathbb{E}(Z) = 0$ et $\mathbb{V}(Z) = 1$.

Remarque :

– La moyenne et l'écart-type d'une variable aléatoire discrète constituent un résumé de la répartition de X. La moyenne donne une idée de l'ordre de grandeur de la variable et l'écart-type une idée de la dispersion des valeurs de X autour de sa moyenne.
– Soit $A \in \mathscr{T}$ et $\mathbb{1}_A$ sa fonction indicatrice. $\mathbb{1}_A$ est une variable aléatoire discrète qui prend la valeur 1 avec la probabilité $P(A)$ et la valeur 0 avec la probabilité $1 - P(A)$. En appliquant la définition, on voit que $\mathbb{V}(\mathbb{1}_A) = P(A)(1 - P(A))$.

Exemple : Un jeu de loto. Le joueur coche 6 numéros sur une grille qui en comporte 49. Les 6 numéros gagnants sont déterminés par tirage au sort. Soit N le nombre de numéros gagnants de la grille. Pour une mise de 2 Euros, on reçoit le gain $G = g(N)$ suivant :

n numéros gagnants	gain $g(n)$	probabilité
6	2 132 885	$7,2 10^{-8}$
5	3575	$7,8 10^{-5}$
4	94	$9,7 10^{-4}$
3	11	$7,8 10^{-2}$

Le gain moyen est $\mathbb{E}(G) = \sum_n g(n)P(N = n) = 1,16$ euros. On peut vérifier que l'écart-type vaut 572. Ainsi le bénéfice moyen du joueur, qui vaut $E(G) - 2 = -0.84$, est négatif, et le jeu est défavorable au joueur. La grande valeur de l'écart-type vient de ce que parfois (mais très rarement), le jeu peut rapporter beaucoup.

d) Propriétés de la variance

- Pour toute constante c, on a $\mathbb{V}(c) = 0$.
- *Formule de Koenig* $\mathbb{V}(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2$
- $\forall a \in \mathbb{R}, \forall b \in /R, \; \mathbb{V}(aX + b) = a^2\mathbb{V}(X)$

e) Variable centrée réduite associée à une variable d'ordre 2

Soit X une variable aléatoire discrète d'ordre 2, m son espérance et σ son écart-type. La variable centrée réduite associée à X est définie par

$$Z = \frac{X - m}{\sigma}.$$

On vérifie facilement en effet que $\mathbb{E}(Z) = 0$ et $\mathbb{V}(Z) = 1$.

4.4 Covariance

Soit $(X, Y) : \Omega \to \mathbb{R}^2$ un couple de variables aléatoires discrètes d'ordre 2, $\sigma^2 = \mathbb{V}(X)$ et $\tau^2 = \mathbb{V}(Y)$.

- La covariance de X et Y est la quantité définie par

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

- On a $\mathbb{V}(X) = cov(X, X)$.Par conséquent, pour tous réels a et b, on a

$$\mathbb{V}(aX + bY) = a^2\mathbb{V}(X) + 2ab\, cov(X, Y) + b^2\mathbb{V}(y)$$

- Le coefficient de corrélation linéaire de X et Y est défini par

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

  On a $-1 \leq \rho(X, Y) \leq 1$. D'autre part, on peut montrer que $|\rho(X, Y)|$ est maximal lorsque Y est une fonction affine de X : $Y = aX + b$. De façon générale, plus le coefficient de corrélation est proche de 1 en valeur absolue, plus les variables X et Y sont linéairement liées. Un coefficient de corrélation nul signifie donc que les deux variables ne sont pas linéairement liées. Il n'empêche qu'elles peuvent être liées par un autre type de relation.

- La matrice de covariance du couple (X, Y) est définie par

$$\begin{pmatrix} \mathbb{V}(X) & cov(X, Y) \\ cov(X, Y) & \mathbb{V}(Y) \end{pmatrix}$$

5 Inégalités

5.1 Inégalité de Tchebychev

Soit X une variable aléatoire discrète de moyenne m et d'écart-type σ alors

$$\forall t > 0,\ P(|X - m| > t) \leq \frac{\sigma^2}{t^2}$$

Interprétation : Si on pose $t = s\,\sigma$, l'inégalité de Tchebychev se réécrit, pour tout $s > 0$

$$\forall s > 0,\ P(|X - m| > s\,\sigma) \leq \frac{1}{s^2}$$

Si on considère l'écart-type σ comme une unité d'écart, ceci dit que la probabilité qu'une variable s'éloigne de plus de s unités d'écart de sa moyenne est inférieure à $1/s^2$.

L'aspect remarquable de l'inégalité de Tchebychev est son universalité, puisqu'elle est vérifiée quelle que soit la loi de la variable X (si tant est bien sûr que celle-ci admette une variance). Le prix à payer est qu'elle ne donne souvent en pratique que de médiocres majorations de la queue de distribution.

5.2 Inégalité de Markov

Soit X une variable aléatoire discrète positive admettant une espérance alors

$$\forall t > 0,\ P(X \geq t) \leq \frac{\mathbb{E}(X)}{t}$$

6 Fonction caractéristique et génératrice

Soit $(X, Y) : \Omega \to \mathbb{R}^2$ un couple de variables aléatoires discrètes, $S = X(\Omega), T = Y(\Omega)$ et $V = S \times T$.

- La fonction caractéristique de (X, Y) est la fonction $\phi : \mathbb{R}^2 \to \mathbb{C}$ définie par

$$\phi(\alpha, \beta) = E(e^{i(\alpha X + \beta Y)}) = \sum_{(s,t)\in V} e^{i(\alpha X + \beta Y)} p_{s,t},\ (\alpha, \beta) \in \mathbb{R}^2$$

  De plus, $\alpha \to \phi(\alpha, 0)$ est la fonction caractéristique de X et $\beta \to \phi(0, \beta)$ celle de Y.
- On suppose que $V \subset \mathbb{N}^2$. La fonction génératrice de (X, Y) est la fonction définie par :

$$g(\alpha, \beta) = E(\alpha^X \beta^Y) = \sum_{i\geq 0, j\geq 0} \alpha^i \beta^j p_{i,j}$$

  De plus, $\alpha \to g(\alpha, 1)$ est la fonction caractéristique de X et $\beta \to g(1, \beta)$ celle de Y.

7 Théorèmes de transfert

7.1 Cas d'une variable aléatoire discrète

Soit $X : \Omega \to \mathbb{R}$ une vad, $S = X(\Omega$ et $h : S \to \mathbb{C}$ une fonction. Alors $h(X)$ est une vad et si $\sum\limits_{s \in S} |h(s)| p_s$ est convergente, la variable $h(X)$ admet une moyenne définie par

$$\mathbb{E}(h(X)) = \sum_{s \in S} h(s) p_s$$

7.2 Cas d'une variable aléatoire discrète

Soit $(X, Y) : \Omega \to \mathbb{R}^2$ un couple de variables aléatoires discrètes, $S = X(\Omega), T = Y(\Omega)$ et $V = S \times T$ et $h : V \to \mathbb{C}$ une fonction. Alors $h(X, Y)$ est une variable aléatoire discrète et si $\sum\limits_{(s,t) \in V} |h(s, t)| p_{s,t}$ est convergente, la variable $h(X, Y)$ admet une moyenne définie par

$$\mathbb{E}(h(X, Y)) = \sum_{(s,t) \in V} h(s, t) p_{s,t}$$

7.3 Théorème de transfert conditionnnel

a) Théorème

Soit $(X, Y) : \Omega \to \mathbb{R}^2$ un couple de variables aléatoires discrètes, $S = X(\Omega)$, $T = Y(\Omega)$ et $V = S \times T$ et $h : V \to \mathbb{C}$ une fonction. Alors $h(X, Y)$ est une variable aléatoire discrète et si $\sum\limits_{(s,t) \in V} |h(s, t)| p_{t|s}$ est convergente, la variable $h(X, Y)$ admet une moyenne par rapport à la lopi conditionnelle de $h(X, Y)$ sachant que $X = s$ définie par

$$\phi(s) = \mathbb{E}(h(X, Y)|X = s) = \sum_{(t) \in T} h(s, t) p_{t|s}$$

Ce résultat permet de calculer la moyenne d'une fonction du couple en passant par le conditionnement.

b) Espérance conditionnelle

L'espérance conditionnelle de $h(X, Y)$ sachant X notée, $E(h(X, Y)|X)$ est définie par

$$E(h(X, Y)|X) = \phi(X)$$

On a alors : $E(h(X, Y)) = E(E(h(X, Y)|X))$

Exemple Le nombre N de voitures passant devant une staion d'essence en un jour suit la loi de Poisson de paramètre $\lambda > 0$. Chaque voiture décide de s'arrêter à la station avec la probabilité p indépendamment des autres. On note K le nombre de voitures qui s'arrêtent à la

station. Que vaut $E(K)$?
Par hypothèse, on a

$$p_N(n) = \frac{\lambda^n}{n!}e^{-\lambda}, p(K|N=n)(k) = C_n^k p^k (1-p)^{n-k}$$

d'où

$$E(K|N=n) = \sum_k kp(K|N=n)(k) = np,$$

soit $E(K|N) = pN$. De plus,

$$E(K) = E(E(K|N)) = E(pN) = pE(N) = p\lambda$$

puisque $E(N) = \lambda$.

Exercices

Exercice 1
Au bord de l'A7, un étudiant fait du stop. En cette saison, un vingtième des automobilistes s'arrête pour prendre un stoppeur. Quelle est la loi du nombre de véhicules que l'étudiant verra passer avant quéil ne trouve un chauffeur ? Quelle est la probabilité qu'il monte dans la quatrième voiture qui passe ? Quelle est la probabilité qu'il voit passer au moins 6 voitures qui ne s'arrêtent pas ?

Exercice 2
Lors d'une visite médicale, n patients de tailles différentes se présentent devant le médecin, et ce dans un ordre aléatoire. On note X le rang de présentation du plus grand d'entre eux. Donner la loi de la variable X.

Exercice 3
On considère un dé cubique truqué, de telle sorte que la probabilité d'obtenir la face numérotée k est proportionnelle à k (on suppose que les faces sont numérotées de 1 à 6). Soit X la variable aléatoire associée au lancer de ce dé.

1. Déterminer la loi de X, calculer son espérance.
2. On pose $Y = 1/X$. Déterminer la loi de Y et son espérance.

Exercice 4
Soit a un entier strictement positif donné et soit X une variable aléatoire discrète dont la loi est donnée par

$$P(X = k) = \frac{1}{a} - \frac{1}{b} \;\; \forall 1 \leq k \leq ab$$

1. Déterminer b
2. Donner la fonction de répartition F de X. La tracer pour $a = 2$.
3. Quelles sont les solutions de l'équation $F(x) = 1/2$?

Exercice 5
Soit X une variable aléatoire de loi de Poisson de paramètre $\lambda > 0$. Calculer $\mathbb{E}((1 + X)^{-1})$.

Exercice 6
On lance deux dés équilibrés. On note $U1$ et $U2$ les variables aléatoires correspondant aux résultats obtenus. On appelle $X = min(U1, U2)$ le minimum et $Y = max(U1, U2)$ le maximum des deux dés.

1. Donner la loi de X. En déduire $E[X]$.
2. Exprimer $X + Y$ en fonction de $U1$ et $U2$. En déduire $E[Y]$.
3. Exprimer XY en fonction de $U1$ et $U2$. En déduire $E[XY]$, puis $Cov(X, Y)$.

Exercice 7
Soit $X \sim \mathcal{G}(p)$ loi géométrique de paramètre $p \in]0, 1[$.

1. Soit $n \in \mathbb{N}$. Exprimer $P(X > n)$ en fonction de p et n. Quel est le lien avec $F(n)$, où F est la fonction de répartition de X ?
2. En déduire la propriété dite "d'absence de mémoire" de la loi géométrique, à savoir que :

$$\forall (m,n) \in \mathbb{N} \times \mathbb{N} P(X > n+m | X > m) = P(X > n).$$

Exercice 8
On jette 3600 fois un dé et on appelle S le nombre de fois où apparaît le numéro 1.

1. Quelle est la loi de S ? Donner sa moyenne et sa variance.
2. Exprimer sous forme d'une somme la probabilité que ce nombre soit compris strictement entre 480 et 720. Grâce à l'inégalité de Tchebychev, minorer cette probabilité.

Exercice 9
Soit $X : \Omega \to \mathbb{N}$ une variable aléatoire de loi de Poisson de paramètre λ. On pose $Y = \dfrac{1}{X+1}$ et $Z = (-1)^X$.
Calculer la moyenne de Y et de Z. Calculer $cov(Y, Z)$.

Exercice 10
La loi du couple (X, Y) est définie par le tableau suivant :

$X \backslash Y$	0	1	2	3
1	$2/48$	$6/48$	$3/48$	$1/48$
2	$4/48$	$10/48$	$6/48$	$2/48$
3	$1/48$	$8/48$	$2/48$	$a/48$

1. Donner la valeur de a.
2. Caculer la loi conditionnelle de Y sachant $X - 2$.
3. On pose $Z = XY$. Calculer $\mathbb{E}(Z)$ et la probabilité que Z soit impaire.
4. Calculer $p = P(Z \geq 6)$. Comparer la valeur de p avec les majorants fournis par les inégalités de Markov et de Tchebychev.

Exercice 11
Un mathématicien se trouve être également un fumeur de pipe et il porte à tout moment deux boîtes d'allumettes, une dans chacune de ses poches gauche et droite. Chaque fois qu'il a besoin d'une allumette, il a une chance sur deux d'aller la chercher dans sa poche gauche et autant pour l'autre. Il découvre subitement que la boîte tirée est vide. Les deux boîtes contenaient au départ N allumettes chacune. Quelle est la probabilité qu'il lui reste k allumettes dans l'autre boîte, $k = 0, 1, ..., N$?
Application numérique : $N = 100, k = 25$. Conclusion ?